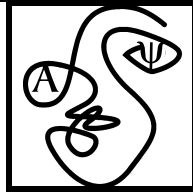# Encyclopedia of
APPLIED
PSYCHOLOGY

Editor-in-Chief

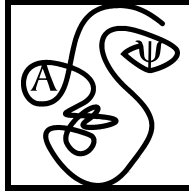# CHARLES SPIELBERGER

VOLUME 1   A–H

Editor–in–Chief

CHARLES D. SPIELBERGER

University of South Florida
Tampa, Florida, USA

Charles D. Spielberger is Distinguished Research Professor of Psychology and Director of the Center for Research in Behavioral Medicine and Health Psychology at the University of South Florida. Author, co-author, or editor of more than 350 professional publications, Spielberger's current research focuses on anxiety, anger, curiosity, depression, job stress, and lifestyle factors that contribute to the etiology and progression of cancer and cardiovascular disorders. His *State-Trait Anxiety Inventory* (1970, 1983), with translations in 66 languages and dialects, has become a standard international measure of anxiety. He has served as President of the International Association of Applied Psychology, the International Council of Psychologists, and the International Stress Management Association and was the 100th President of the American Psychological Association.

Editorial Advisors

Section Editors

Preface

During the past 50 years, there has been remarkable progress in the articulation of psychological theory, increasingly sophisticated reports of empirical research, and a dramatic expansion in professional applications of psychology in all areas of modern life. Applied psychology requires both a sound scientific foundation and attention to the practical problems that are encountered in applying psychology to an infinite variety of life settings. Growing public recognition of the potential contributions of psychology to solving life's problems has also fostered rapid growth in the number of psychologists who are engaged in research and in professional practice throughout the world.

The need for an authoritative publication that reviews and evaluates the numerous applications of professional psychology has been recognized for a number of years, and has been of special interest to the members of the International Association of Applied Psychology (IAAP), the world's oldest and largest membership organization of individual psychologists. The interests and concerns of IAAP members, who reside in more than 80 countries, have been strongly noted by former IAAP presidents Edwin A. Fleishman, Claude Levy-Leboyer, and Harry C. Triandis.

Significant contributions to international applications of psychology were explicitly recognized and discussed in the context of their diverse societal settings by Bernhard Wilpert in his Presidential address at the IAAP's 24th International Congress of Applied Psychology in 1998. It was my honor and pleasure to succeed Professor Wilpert as President of the IAAP at this Congress, which was held in San Francisco, California, and to begin participating in discussions to identify the most important contributions to applied psychology that are reported in this encyclopedia.

The development and current status of applied psychology can be best understood by examining the historical evolution of the field, which is cogently described by Helio Carpintero, the official Historian of the IAAP. In his overview article on the history of applied psychology, Carpintero highlights the benefits resulting from psychologists joining together to form professional associations that have facilitated the development and integration of psychological theory and research and that have stimulated and guided more effective applications of psychology in professional work.

Founded in 1920 as the *Association Internationale des Conferences de Psychotechnique* (International Association for Psychotechnology), the IAAP adopted its present name in 1955. Edouard Claprēde convened and chaired the first meeting of the association, held in Geneva, Switzerland in 1920, and served as its President for more than 20 years. Meetings were held in a number of European cities, including Barcelona, Milan, Paris, Moscow, and Prague. An excellent history of the evolution and development of applied psychology and the IAAP can be found in the address given by former President Edwin A. Fleishman on receiving the 1999 American Psychological Association's Award for Distinguished Contributions to the International Advancement of Psychology (*American Psychologist*, November, 1999, pp. 1008–1016).

For recommending that the IAAP take an active role in sponsoring and developing the *Encyclopedia of Applied Psychology*, we are indebted to George Zimmar, an active IAAP member and former senior sponsoring editor at Academic Press, now Elsevier. The first step in developing the encyclopedia was the appointment of an Advisory Board composed of distinguished colleagues representing all areas of

applied psychology. In addition to recommending the major areas and the specific topics to be covered in the encyclopedia, the Advisory Board identified highly qualified colleagues who would be invited to serve as Section Editors for the selected areas.

The responsibilities of the Section Editors included writing a major article to provide an overview and analysis of their respective areas and working with the Editor-in-Chief and the Advisory Board in identifying the specific topics to be included in each area. They also helped to determine the amount of space to be allocated for each topic and recommended colleagues with outstanding knowledge and expertise to serve as the authors for the individual articles. In addition, the Section Editors reviewed and edited the articles that presented information on the topics that were related to the subject matter that was included in their sections.

The articles that are included in the encyclopedia are presented in alphabetical order, as determined by the title of each article. Consequently, the underlying section organization will not be readily apparent. At the conclusion of each article, the reader is directed to additional articles that include related information. The authors of each article also provide a number of references that are recommended for further reading.

It should be noted, however, that it was not possible to cover all of the many areas of applied psychology. Several emerging areas are not included because we failed to recognize their importance at the time the topics were selected for the encyclopedia. It is also unfortunate that a few authors were unable to meet the deadlines for submitting the manuscripts for their articles.

The *Encyclopedia of Applied Psychology* is truly a product of the collaborative efforts of the members of the IAAP. We are especially indebted to former IAAP Presidents Edwin A. Fleishman, Harry C. Triandis, and Bernhard Wilpert, and to our current president, Michael Frese, for their leadership, encouragement, and tangible contributions to the encyclopedia. For their dedicated contributions to the encyclopedia, I would also like to thank the members of the Advisory Board and the Section Editors, whose names are listed in the front matter. Finally, on behalf of the IAAP, I thank the authors who contributed articles in their areas of expertise and the editorial and production staff of Elsevier for the tremendous amount of time and energy that was needed to bring this large project to fruition.

CHARLES D. SPIELBERGER, Ph.D., ABPP
President, IAAP, 1998–2002

A

Academic Failure, Prevention of

Mary Ann Rafoth

Indiana University of Pennsylvania, Indiana, Pennsylvania, USA

GLOSSARY

basic skills The skills associated with functional literacy, including reading decoding and comprehension skills, math calculation and problem solving, and writing.

cooperative learning An instructional approach that uses peer groups to facilitate and reinforce learning.

curriculum-based assessment An assessment approach that uses students' progress in their actual curriculum as a measurement point.

emergent literacy The recognition of environmental print that leads to gradual understanding of the role of letters and words in written language.

learning disability A discrepancy between a student's measured ability and actual achievement caused by a dysfunction in a basic neuropsychological process.

learning strategy A voluntary activity students use to facilitate remembering, learning, or problem solving.

mastery learning An instructional approach that allows students to repeat instruction and assessment until competency is achieved.

peer tutoring An instructional approach that uses peers to reinforce classroom instruction by providing individual instruction to students.

phonological awareness Children's awareness of the way in which words sound, particularly rhyme and alliteration and the ability to blend sounds, recognize onset rime (initial sounds), and identify sound units.

retention The practice of requiring a student to repeat a grade or requiring a child of appropriate chronological age to delay entry to kindergarten or first grade.

school readiness The set of skills children need to learn successfully in school.

social promotion The practice of promoting a student to the next grade even when skills or knowledge sets to be mastered at that level have not been mastered.

team teaching An instructional approach that allows teachers, often with different areas of expertise, to share instructional duties.

Prevention of academic failure is a serious challenge because children who fail academically experience significant social and economic challenges throughout their lives. Causes of academic failure include familial, socioeconomic, and cultural issues that lead to a lack of readiness for school, academic, instructional, and motivational problems as well as physiological, cognitive, and neurological barriers to learning. Attempts to help students who are experiencing academic failure fall into three categories: prevention, intervention, and remediation. Preventive approaches

aim to stop academic failure before it occurs. Early intervention programs aim to catch children during key developmental periods and facilitate development and readiness skills. Remediation programs are usually applied when students have demonstrated significant skill deficits and are experiencing significant academic failure. Special education programs often take this form, as do other kinds of academic accommodations for students identified with special needs.

1. DEFINING ACADEMIC FAILURE

A lot depends on children's success in school—their self-esteem, their sense of identity, their future employability. Preventing academic failure means that we, as a society, are much more likely to produce individuals who feel confident about their ability to contribute to the common good, whose literacy skills are competent, and who are able to hold jobs successfully. Thus, prevention of academic failure should be a primary concern for any society. But exactly what is meant by academic failure? What does the term connote? Generations of schoolchildren since the 1920s, when the system of grade progression began, have equated academic failure with retention in grade. School failure meant literally failing to progress onto the next grade, with the assumption that the skills and knowledge taught in that grade had not been mastered. To have flunked multiple grades quickly led to quitting school altogether—the ultimate academic failure.

More recently, academic failure has come to mean a failure to acquire the basic skills of literacy. Students who were unable to read at a functional level, to communicate effectively through writing, and to complete basic math calculations were seen as representing a failure of the academic system even though they might hold high school diplomas. The practice of moving students on from one grade to the next even though they might not have mastered basic competencies associated with lower grade levels is often referred to as social promotion. This type of academic failure led to calls for an increased emphasis on basic skills, that is, the "three R's"—reading, (w)riting, and (a)rithmetic—in public education. Partly in reaction to emphasis on basic skills, a third interpretation of academic failure has also emerged. In this view, academic failure occurs not only when students fail to master basic skills but also when they emerge from school without the ability to think critically, problem solve,

learn independently, and work collaboratively with others—a skill set deemed necessary for success in a digital age. This underachievement symbolizes a significant loss of intellectual capital for a culture. Finally, statistics show that students who do not complete high school are much more likely to need welfare support, have difficulties with the law and police, and struggle economically and socially throughout their lives. Thus, academic failure ultimately means both the failure to acquire the skill sets expected to be learned and the failure to acquire official documentation of achievement by the school system.

2. CAUSES OF ACADEMIC FAILURE

Students struggle academically for many reasons, including familial, socioeconomic, and cultural issues that lead to a lack of readiness for school, academic, instructional, and motivational problems as well as physiological, cognitive, and neurological barriers to learning. Early school failure often occurs because children enter the structured school environment not ready to learn.

2.1. School Readiness

School readiness refers to the idea that children need a certain set of skills to learn and work successfully in school. Often this term refers to whether or not children have reached the necessary emotional, behavioral, and cognitive maturity to start school in addition to how well they would adapt to the classroom environment. To create some consensus about when a child should begin school, states designate a specific cutoff date. If a child reaches a certain age by the cutoff (usually 5 years for kindergarten and 6 years for first grade), the child may begin school. However, cutoff dates are arbitrary and vary considerably across nations, and age is not the best determinant or most accurate measure of whether or not a child is ready to begin school. Research has suggested that we must look at all aspects of children's lives—their cognitive, social, emotional, and motor development—to get an accurate idea of their readiness to enter school. Most important, children's readiness for school is affected by their early home, parental, and preschool experiences.

Stated in its simplest form, school readiness means that a child is ready to enter a social environment that is focused primarily on education. The following list of

behaviors and characteristics are often associated with school readiness:

- Ability to follow structured daily routines
- Ability to dress independently
- Ability to work independently with supervision
- Ability to listen and pay attention to what someone else is saying
- Ability to get along with and cooperate with other children
- Ability to play with other children
- Ability to follow simple rules
- Ability to work with puzzles, scissors, coloring, paints, and the like
- Ability to write own name or to acquire the skill with instruction
- Ability to count or acquire skills with instruction
- Ability to recite the alphabet
- Ability to identify both shapes and colors
- Ability to identify sound units in words and to recognize rhyme

Family environment is very important in shaping children's early development. Some family factors that can influence school readiness include low family economic risk (poor readiness for school is associated with poverty), stable family structure (children from stable two-parent homes tend to have stronger school readiness than do children from one-parent homes and from homes where caregivers change frequently), and enriched home environment (children from homes where parents talk with their children, engage them in conversation, read to them, and engage in forms of discipline such as "time-out" that encourage self-discipline have stronger readiness skills).

Children's readiness to read, in particular, has gained greater attention from educators recently as the developmental precursors to reading have become more evident. During the preschool years, children develop emerging literacy skills—preacademic skills that allow children to develop a disposition to read, write, and compute. Children are ready to read when they have developed an ear for the way in which words sound and can identify rhyme and alliteration, blend sounds, recognize onset rime (initial sounds), and identify sound units in words. Together, these skills are called phonological awareness and usually emerge in children between 2 and 6 years of age. Children with good phonological awareness skills usually learn to read quickly. Children who are poor readers have weak phonological skills, and children who do not learn to read fail in school. Another important readiness skill that helps children

to learn to read is called print awareness. Print awareness means that children are capable of the following:

- Knowing the difference between pictures and print
- Recognizing environmental print (e.g., stop signs, McDonald's, Kmart)
- Understanding that print can appear alone or with pictures
- Recognizing that print occurs in different media (e.g., pencil, crayon, ink)
- Recognizing that print occurs on different surfaces (e.g., paper, computer screen, billboard)
- Understanding that words are read right to left
- Understanding that the lines of text are read from top to bottom
- Understanding the function of white space between words
- Understanding that the print corresponds to speech word for word
- Knowing the difference between letters and words

Children also need to learn book-handling skills such as orienting a book correctly and recognizing the beginning and the end of a book. Children who begin school without these basic readiness skills are at risk for school failure. The use of screening assessments during preschool and kindergarten to identify students who may be at risk for academic failure, particularly in the area of phonemic awareness, has been shown to be a sound method of predicting which children will have difficulty in learning to read. Most likely to be retained in kindergarten are children who are chronologically young for their grade, developmentally delayed, and/or living in poverty.

2.2. Academic, Instructional, and Motivational Reasons

Children who do not master basic reading skills, specifically the ability to automatically decode new words and build a sight word vocabulary that leads to fluency, experience academic failure. By third grade, learning to read has become reading to learn. In other words, in third grade the curriculum becomes focused much less on teaching students to acquire the basic tools of literacy (reading, writing, and computing) and much more on using those tools to learn content, express ideas, and solve problems. At this point, students are likely to be given content textbooks in science and social studies and to read nonfiction for the purpose of gaining new information. Thus, the inability to read effectively and to learn

to study independently often leads to failure at the elementary and middle school levels and also creates profound motivation problems at the high school level that contribute to the ultimate school failure—dropping out.

The inability to master key concepts in pivotal classes such as algebra, now typically taken at the middle or junior high school level, often limits students' ability to proceed in coursework. Students may fail to understand algebraic concepts due to their developmental level. (Many students are stilling thinking in concrete terms in middle school and have not yet moved into a stage of cognitive thinking allowing them to understand formal logic and manipulate symbols—a developmental source of failure.) In addition, some students might not have automatized basic arithmetic skills, particularly computing with fractions—an academic or instructional failure. Some students may have become turned off to math and accepted self-images that permit poor math skills—a motivational failure. Finally, many students will fail algebra for all of these reasons, and the impact will often be that they will finish school in a nonacademic or basic track or might even drop out.

Thus, academic and instructional reasons for school failure include the effectiveness of the instruction a student has received and the quality of remediation strategies or programs available. The following is a typical example that illustrates academic and instructional reasons for school failure. A teacher reports that a student is having difficulty in getting beyond the primer level in reading and is being considered for retention. The child was assessed as having average intelligence. No behavioral or attention problems were noted. Closer inspections of the student's reading skills indicated that she had poor phonological skills and was not profiting from the type of classroom reading instruction she was receiving that depended heavily on auditory phonics instruction stressing "sounding out words" and matching sound–symbol connections. Appropriate interventions included using techniques to build up a sight word vocabulary through repetition and distributed learning and introducing the student to a visual decoding system to provide her with a method for reading unknown words by analyzing the words and breaking them down into more familiar visual units.

2.3. Physiological, Neurological, and Cognitive Reasons

Imagine a child spending most of the year in kindergarten with an undetected hearing loss that has made it very difficult for her to benefit from instruction. Imagine another child in first grade struggling to learn because her vision impairment has not been caught or corrected. Similarly, students suffering from a variety of conditions and illnesses, such as childhood diabetes, asthma and allergy-related problems, and sickle cell anemia, may have difficulty in maintaining energy and attention in school due to chronic fatigue and the impact of medications. Children may also suffer from orthopedic or motor impairments that make it difficult for them to explore their environment, interact with others, and/or master tasks that demand motor skills.

Students who suffer from various kinds of neurological disorders or learning disabilities may also have cognitive learning problems that make it difficult for their brains to process information, interpret sounds and symbols efficiently in reading, calculate and understand number concepts, and/or write effectively. Other children may have cognitive deficits, such as mental retardation, that limit their ability to absorb and apply regular classroom instruction. Children with attention deficit disorders have difficulty in directing and maintaining their attention, may exhibit impulsive behavior, and have trouble in interacting independently in typical classroom environments without support. Specialized and/or special education interventions are designed to provide individualized strategies and approaches for students who have physiological-based learning problems interfering with their ability to learn.

3. PREVENTING EARLY SCHOOL FAILURE

3.1. Early Intervention Programs

Programmatic interventions may include developing screening programs to identify children at risk for school failure and to ensure early access to readiness programs already available in the school or community such as Head Start. Many states are now developing guidelines for children age 6 years or under based on the National Association for the Education of Young Children's (NAEYC) list of developmentally appropriate practices. The major challenge facing early intervention programs is to provide developmentally and individually appropriate learning environments for all children. Essential ingredients to successful preschool experiences include small group and individualized teacher-directed activities as well as child-initiated

activities. Quality programs recognize the importance of play and view teachers as facilitators of learning.

3.2. Preventing School Failure in the Elementary Grades

Full day kindergarten (as opposed to half day) programs provide more time for field trips, activity centers, projects, and free play. At-risk students who attend rigorous yet nurturing full day programs have a greater chance of experiencing academic success. Full day kindergarten programs help increase academic achievement as well as decrease the number of children retained in the early elementary grades. Research shows that full day kindergarten programs for children who come from disadvantaged backgrounds lead to stronger achievement in basic skill areas and generally better preparation for first grade.

Table I summarizes the research on full-day kindergarten. Any decisions about whether or not to schedule full- or half-day programs should recognize that what a child is doing during the kindergarten day is more important than the length of the school day.

The instructional technology that enables classroom teachers to meet the needs of students of different skill levels is already available, but in many cases teachers do not have access to that technology. Reading interventions that provide intensive, early, and individualized help that targets a child's specific weaknesses (e.g., Success for All, Reading Recovery, Direct Instruction) have been shown to be effective in reducing early reading failure. Instructional approaches such as mastery learning, adaptive education, team teaching, cooperative learning, peer tutoring, and curriculum-based assessment are methods that have been shown to produce academic gains in students of all achievement levels in the elementary grades. Recently, technology has offered greater individualization of instruction and increased flexibility in allowing students to progress at their own pace and to respond to instruction.

TABLE I
Research Summary: Full-Day Versus Half-Day Kindergarten Programming

Research on effects of full-day programs on achievement	*Research on effects of full-day programs with educationally disadvantaged and low-SES children*	*Research on other effects of full-day versus half-day programs*
Full-day programs increase long-term achievement.	Full-day programs allow for greater opportunity for assessment and screening.	Full-day programs have been shown to raise children's self-esteem, independence, and creativity. There is no evidence that the curriculum is more individualized or innovative.
Full-day programs produce higher reading scores in second and third grades.	Full-day programs result in higher academic and social gains.	Parental involvement is diminished in full-day programs.
Full-day programs result in fewer grade retentions, higher report card marks, and higher scores on standardized tests.	Full-day programs allow for more informal and formal instructional time on an individual basis.	Full-day programs are not found to cause excessive fatigue or stress in children.
Full-day programs provide more time for individualized instruction, academics, and the reinforcement of children's positive self-images.	Full-day programs provide access to nutritional lunch as well as breakfast.	Full-day programs allow for an unhurried relaxed school day with more variety of experiences.
Full-day programs are found to have no effect on achievement.	Title I children have significantly higher achievement in full-day programs.	Children in full-day programs have been shown to be less likely to be dependent, shy, and withdrawn. Children in half-day programs spend more time in large groups than in free play.

Note. SES, socioeconomic status.

4. PREVENTING FAILURE IN THE INTERMEDIATE GRADES AND MIDDLE SCHOOL

Remedial programs, such as the Title I or Chapter 1 programs, have also been used to remediate early skill deficits in reading and math. However, developing intervention programs, such as after-school tutoring and summer school courses, might not be sufficient to make up serious deficits in short amounts of time and cannot take the place of preventive systemic approaches.

The use of learning strategies instruction has been shown to be very effective in improving study skills and performance in middle school students. Because unsuccessful middle school students often lack basic strategic learning skills, intervention programs should also target these areas. Similarly, approaches that use learning, problem-solving, and memory strategies are the most effective interventions in terms of producing actual gains in student achievement in the classroom.

5. PREVENTING ACADEMIC FAILURE IN HIGH SCHOOL

At the secondary level, development of reentry programs for dropouts and alternative education programs, such as those that combine teaching skills with job training, are essential to prevent further academic failure. Research on academic failure at the secondary level has generally examined the relationships between grade retention and attendance, suspension, and self-concept, with an emphasis on the correlation between retention and dropout rates. Academic failure at the high school level is related to attendance and suspension rates. In general, students who are failing do not attend school on a regular basis. In addition, students who have been retained prior to the secondary level are less likely to attend school on a regular basis in junior and senior high school. Furthermore, regardless of the grade in which retention occurs, secondary students who have been retained often exhibit low self-esteem.

Many studies have reported that students who drop out are five times more likely to have repeated a grade than are students who eventually graduate. Being retained twice virtually guarantees that a student will drop out of school, and grade retention alone has been identified as the single most powerful predictor of dropping out. The dropout rate of overage students is appreciably higher than the dropout rate of regularly promoted students when reading achievement scores are equivalent for the two groups. Even in high-socioeconomic school districts, where students are less likely to leave school, a significant increase in dropout rates has been found for retained students.

Successful programs at the high school level often have two characteristics: (a) one or more individuals who develop relationships with students individually and monitor their progress carefully and (b) some mechanism to allow students who have failed courses and lost credits to regain these credits in quicker than normal time, allowing for graduation at the expected time. Simply put, successful programs must address the motivational issues that have developed by adolescence and the lack of academic achievement identity typically present in students who drop out of school. School-to-work programs that combine vocational counseling with on-the-job experience are successful ways in which to increase a sense of academic competence while connecting to students' current self-concepts and needs.

6. CONCLUSION

Attempts to help students who are experiencing academic failure fall into three categories: prevention, intervention, and remediation. Preventive approaches aim to stop academic failure before it occurs. Early intervention programs from birth to 5 years of age, for example, aim to catch children during key developmental periods and facilitate development and readiness skills. Intervention programs, such as Robert Slavin's Success for All program, aim to intervene as soon as students begin to show signs of slipping behind their peers. Intervention plans may also be designed under Section 504 of the Americans with Disabilities Act, which mandates accommodations in the instructional environment for students who have physical or neurological problems that may interfere with their ability to learn or succeed in a typical classroom. Remediation programs are usually applied when students have demonstrated significant skill deficits and are experiencing significant academic failure. Special education programs often take this form, as do other kinds of academic accommodations for students identified with special needs.

Of course, early identification and prevention of academic problems is always preferable to later intervention and remediation. Thus, systemic solutions that target early reading deficits, independent learning skills, and motivational problems from a developmental

perspective are essential to the prevention of academic failure. Working to change school practices will require sharing the research with educators, conducting evaluations on the outcomes of alternative interventions at the local level, and lobbying at the state level to promote changes in policy and to advocate for alternative service delivery systems that more effectively meet the needs of students experiencing school failure. Successful programs to boost student achievement, however, must attack underachievement in three key areas. These key areas—early reading intervention, acquisition of strategic learning and study skills, and motivation to achieve—are highly related to school failure.

First, acquisition of basic reading skills must be addressed. If students underachieve in the primary grades, it is most often because they have failed to learn to read. Kindergarten screenings should include an assessment of phonological awareness. Children identified with weak skills should be targeted for intervention through phonological awareness training in kindergarten. Prekindergarten programs for high-risk students are recommended. Students should be tracked using curriculum-based assessments of oral reading in the primary grades. Any student who falls behind the average rate of acquisition for his or her class should receive an individualized analysis of reading skill and additional after-school intervention based on that analysis to allow the student to "catch up" to classmates. This early, intensive, and individualized intervention allows for all students to enter the intermediate grades as able readers. Some students with special needs might not progress at the same rate as their classmates, but they too will benefit from early reading interventions.

Second, students must acquire independent learning and study skills during the intermediate and middle school years if they are to maximize achievement and be competitive in the job market of tomorrow. Many students underachieve in middle school because they lack the organizational and learning strategies to master the demands of the upper grades. Embedded approaches to strategy instruction facilitate generalization and encourage students to use all of their mental tools. Assessment of students' study skills and metacognitive development (i.e., the degree to which they are aware of and control their own cognitive processes) leads directly to specific interventions.

Third, students in high school often underachieve because they lack the motivation to excel academically. They often have failed to incorporate pictures of themselves as successful students into their self-concepts. Through a variety of approaches, including staff inservice, a study skills coach approach to peer tutoring, and an individualized profile of each student's study style and vocational options, increased academic competence and a value for academic work can be built.

See Also the Following Articles

Academic Interventions ▪ Advertising and Culture ▪ Educational Achievement and Culture ▪ Educational and Child Assessment ▪ Effective Classroom Instruction ▪ Learning ▪ Learning Styles and Approaches to Studying ▪ Mathematics, Teaching of ▪ Reading, Teaching of ▪ Teaching Effectiveness ▪ Transfer of Learning ▪ Writing, Teaching of

Further Reading

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn*. Washington, DC: National Academy of the Sciences Press.

Elicker, J. (2000). *Full day kindergarten: Exploring the research*. Bloomington, IN: Phi Delta Kappa International.

Gredler, G. (1992). *School readiness: Assessment and educational issues*. Brandon, VT: CPPC Publishing.

Jimerson, S. R. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37, 243–272.

Minke, K. M., & Bear, G. G. (Eds.). (2000). *Preventing school problems, promoting school success: Strategies and programs that work*. Silver Spring, MD: National Association of School Psychologists.

National Institute of Child Health and Development. (2000). *Report of the National Reading Panel: Report of the subgroups*. Bethesda, MD: Author.

Slavin, R. E., Karweit, N. L., & Madden, N. A. (1989). *Effective programs for students at risk*. Boston: Allyn & Bacon.

Smith, M. L., & Shepard, L. A. (Eds.). (1989). *Flunking grades: Research and policies on retention*. New York: Falmer.

Springfield, S., & Land, D. (Eds.). (2002). *Educating at-risk students*. Chicago: National Survey of Student Engagement. (Distributed by University of Chicago Press).

Academic Interventions

Natalie Rathvon

Center City Consortium Schools, Washington, D.C., USA

This article reviews the field of academic interventions. It describes types and targets of academic interventions; intervention delivery systems; guidelines relating to the selection, implementation, and evaluation of academic interventions; and selected evidence-based strategies.

GLOSSARY

academic enablers Nonacademic skills, behaviors, and attitudes that permit students to benefit from instruction.

academic engaged time The proportion of instructional time in which students are actively engaged in learning.

academic intervention A modification in the classroom environment and/or instructional practices designed to produce positive changes in academic performance.

intervention assistance programs (IAPs) Programs based on a consultation model of service delivery and designed to help teachers provide interventions in the regular classroom for difficult-to-teach students; they are also called prereferral intervention programs.

intervention scripts Detailed guidelines to assist intervention agents in delivering an intervention as intended.

social validity The degree to which an intervention is judged as acceptable and important by intervention participants and stakeholders, including the social significance of the goals, the social acceptability of the procedures, and the social importance of the outcomes.

treatment integrity The degree to which an intervention is implemented as planned.

1. OVERVIEW

Today's educators are encountering increasing numbers and proportions of students who have trouble achieving at grade-level expectations. Whether these students are termed difficult to teach, learning disabled, or at risk, they fail to respond to traditional instructional methods. The need for academic interventions is underscored by reports from the National Assessment of Educational Progress (NAEP) indicating that fewer than one-third of American fourth- and eighth-grade students are performing at proficient levels in reading, mathematics, and writing. Recent federal legislation, including the 1997 reauthorization of the Individuals with Disabilities Education Act, which promotes the use of interventions prior to referral for special education services, and the reauthorization of the Elementary and Secondary Education Act in the No Child Left Behind Act of 2001, with its mandates for empirically validated practices, is also spurring interest in strategies that can enhance teachers' capacity to meet students' needs and students' capacity to respond to instruction.

2. TYPES OF ACADEMIC PROBLEMS

Academic problems may be characterized as skill deficits, fluency deficits, performance deficits, or some combination of these. Skill deficits refer to deficiencies reflecting inadequate mastery of previously taught academic skills. Fluency deficits refer to deficiencies in the rate at which skills are performed accurately. Students with performance deficits possess adequate skills and fluency but do not produce work of satisfactory quantity, quality, or both. Many of the interventions presented in Table I are designed to enhance skill acquisition and fluency by increasing opportunities to respond to academic material. Others target performance problems by using self-monitoring or contingency-based procedures, especially group-oriented contingencies that capitalize on peer influence to encourage academic productivity and motivation.

3. INTERVENTION DELIVERY SYSTEMS

Academic interventions can be implemented through a variety of delivery systems, including (a) case-centered teacher consultation, (b) small-group or classroom-centered teacher consultation, (c) staff development programs, and (d) intervention assistance programs (IAPs). For academic interventions with a home component, intervention services can be delivered through case-centered parent consultation, parent training programs, or parent participation in IAPs.

3.1. Intervention Assistance Programs

IAPs are based on a consultation model of service delivery and are designed to increase the success of difficult-to-teach students in the regular classroom by providing consultative assistance to teachers. Since the 1990s, IAPs have become widespread, with the majority of states now requiring or recommending interventions prior to special education referral. Several IAP approaches based on collaborative consultative models of service delivery have been developed to meet the needs of difficult-to-teach students in the regular classroom. These models fall into two general categories depending on whether special education personnel are involved. Key factors in successful implementation and maintenance of IAPs include administrative support, provision of high-quality

interventions, and support of teachers during the intervention process. Although an increasing body of evidence supports the efficacy of IAPs in reducing referrals to special education and improving teachers' attitudes toward diverse learners, relatively few studies have documented that IAPs produce measurable gains in student performance, and many of the studies reporting academic improvement suffer from methodological problems.

4. GUIDELINES FOR SELECTING, IMPLEMENTING, AND EVALUATING ACADEMIC INTERVENTIONS

In designing interventions for students with academic problems, intervention effectiveness can be enhanced by following nine guidelines that are presented in Table II and serve as the foundation for the discussion that follows. The guidelines reflect the importance of balancing treatment efficacy with usability considerations to accommodate the realities of today's classrooms.

4.1. Selecting Academic Interventions

Despite the growing database of evidence-based interventions, studies indicate that teachers and IAPs continue to rely on interventions characterized by familiarity or ease of implementation rather than on those with documented effectiveness. The importance of considering efficacy in intervention selection has been underscored by the Task Force on Evidence-Based Interventions in School Psychology, jointly sponsored by the Division of School Psychology of the American Psychological Association and the Society for the Study of School Psychology and endorsed by the National Association of School Psychologists. Founded in 1998 as an effort to bridge the often-cited gap between research and practice, the Task Force has developed a framework with specific efficacy criteria for evaluating empirically supported intervention and prevention programs described in the literature.

Consultants should also give priority to proactive interventions that help teachers to create learning environments that prevent academic problems from occurring by promoting on-task behavior and productivity rather than using reactive strategies that are applied after problems have already developed. Academic achievement is significantly related to the amount of time allotted for instruction and to academic engagement

TABLE I
Descriptions of Selected Academic Interventions[a]

Intervention	Treatment components	Description	Target(s)	Primary intervention agent(s)
Interventions to improve academic productivity				
Self-management	Goal setting Self-monitoring Self-evaluation Self-reinforcement	Students set goals for work completion and/or accuracy and monitor their own progress	On-task behavior Classwork completion and/or accuracy Homework completion and/or accuracy Academic achievement	Students
Classwide peer tutoring	Increased opportunities to respond Timing Self-management Peer-mediated instruction and feedback Error correction	Students practice oral reading, math facts, or other basic academic skills in pairs	On-task behavior Classwork completion and/or accuracy Academic achievement	Peers
Interventions to improve reading performance				
Listening previewing	Modeling Guided practice Performance feedback Error correction	Students read a passage aloud after listening to a proficient reader read the same passage	Reading accuracy and fluency	Teacher, peers, or parents
Collaborative strategic reading	Strategic instruction Modeling Guided practice and feedback Peer-mediated instruction and feedback	Students learn four basic reading strategies and then work in pairs or groups to implement the strategies	Reading comprehension and vocabulary Study skills	Teacher and peers
Interventions to improve mathematics performance				
Cover–copy–compare	Increased opportunities to respond Error correction Timing Self-instruction, evaluation, and reinforcement	Students respond to questions, compare their answers with instructional prompts, and correct their own errors	Academic productivity, accuracy, and fluency	Teacher

Continues

Intervention	Treatment components	Description	Target(s)	Primary intervention agent(s)
Reciprocal peer tutoring	Increased opportunities to respond Error correction Timing Self-management Peer-mediated instruction and feedback Home-based reinforcement (optional)	Tutoring pairs select their own math goals and practice computational skills in pairs	Mathematics productivity, accuracy, and fluency	Peers
Interventions to improve written language performance				
Self-regulated strategy development in writing	Explicit teaching Modeling Guided feedback Goal setting Self-monitoring, self-instruction, and self-reinforcement Collaborative practice	In a series of lessons, students learn writing skills and strategies as well as procedures for regulating strategy use and the writing process	Writing mechanics Written expression skills Attitudes toward writing	Teacher Peers
Self-recording of writing productivity	Goal setting Self-monitoring Self-evaluation Individual or group reinforcement (optional)	Students evaluate their progress in writing by counting and recording the numbers of words they write	Writing fluency Writing productivity Attitudes toward writing	Students
Interventions to promote academic enablers				
Public posting	Performance Feedback Teacher praise Individual or group contingencies (optional)	Student grades or work completion rates are displayed individually or by teams	On-task behavior Classwork completion and/or accuracy Homework completion and/or accuracy	Teacher
School-home notes	Performance feedback Teacher and parent praise Home-based contingencies (optional)	Students share ratings of academic productivity, and/or achievement with parents and may receive social or material rewards for improvement	Classroom behavior Classwork and/or homework accuracy and/or productivity Attitudes toward school	Teacher Parent

[a]Adapted from *Evidence-Based School Interventions* by N. Rathvon (in press) New York: Guilford Press. Adapted with permission.

A. Select research-based strategies with a documented history of success with the target student population, target setting, and target skill or subskill.

B. Give priority to strategies that are proactive in nature and have the potential of enhancing the achievement of groups of students.

C. Select interventions that are minimally intrusive, fit into the ecology of the classroom setting, and capitalize on the human and material resources already present in school ecosystems.

D. Involve key stakeholders, including teachers, parents, and students themselves, if developmentally appropriate, in the intervention process.

E. Be sensitive to cultural diversity issues in the consultation and intervention processes.

F. Use a variety of training formats to expand teachers' knowledge of interventions and the intervention process.

G. Build in treatment integrity checks to ensure accurate implementation.

H. Provide classroom-based assistance to teachers in the form of modeling, guided practice, and feedback during the implementation and evaluation phases.

I. Design evaluation strategies for assessing changes in student performance that are reliable, valid, and practical for regular classroom teachers.

[a]Adapted from *Effective School Interventions: Strategies for Enhancing Academic Achievement and Social Competence* by N. Rathvon, 1999, New York: Guilford Press. Copyright 1999 by Guilford Press. Adapted with permission.

rates, that is, the proportion of instructional time in which students are actively engaged in learning as demonstrated by behaviors such as paying attention, working on assignments, and participating in class discussions. Although all students profit from proactive strategies that increase instructional time and academic engagement, such interventions are especially important for diverse learners, who are more likely to need additional practice on academic tasks to keep up with their grade peers.

Academic interventions should be minimally intrusive so that they can be implemented in regular classroom settings without unduly disrupting instructional and management routines. Interventions with low ecological validity (i.e., strategies that require major alterations in classroom procedures or cumbersome reinforcement delivery systems) are unlikely to become integrated into regular education routines. Priority should also be given

to strategies that benefit more than one student. Traditional case-centered intervention approaches directed at a single low-performing student are of limited utility in enhancing teachers' overall instructional effectiveness and can be very time-consuming. Group-focused strategies are both time- and labor-efficient, are more acceptable to teachers and students alike, and foster a positive classroom climate by incorporating peer influence.

Moreover, intervention targets and procedures should be reviewed in terms of their social validity. Although many interventions target on-task behavior, increasing on-task behavior is less socially significant as an intervention goal than is enhancing rates of academic responding because increasing on-task behavior does not necessarily result in higher student achievement. Although there is little systematic research on the benefits of involving parents and students in designing interventions, input should be obtained from all stakeholders on the acceptability of proposed intervention goals and procedures.

In designing interventions, consultants must be sensitive not only to individual differences but also to differences in cultural values and norms. For example, interventions that provide concrete operant reinforcers for academic performance might be considered unacceptable by individuals from certain cultures. Unfortunately, little research to date has examined the relative efficacy and acceptability of various academic interventions with culturally and linguistically diverse learners. However, by maintaining an awareness of their own ethnocentrism and encouraging an open dialogue throughout the intervention process, consultants will be better prepared to adjust intervention services to address the needs of stakeholders from nonmainstream groups.

4.2. Implementing Academic Interventions

Regardless of the quality of the intervention design and the documented efficacy of the intervention components, no strategy will be effective in improving student achievement unless the teacher implements it accurately and consistently, that is, with treatment integrity. In the absence of treatment integrity measures, there is no way to determine whether changes in academic performance are due to the effects of the intervention or to factors that are unrelated to intervention components. Among the factors influencing treatment integrity are (a) intervention complexity, (b) time and

material resources required for implementation, (c) the number of intervention agents, (d) efficacy (actual and as perceived by the intervention agents and stakeholders), and (e) the motivation of the intervention agents and stakeholders. Strategies for enhancing treatment integrity include (a) delivering interventions by means of a videotape or an audiotape, (b) documenting consultation contacts, (c) using an intervention manual or script, (d) having a written intervention plan, and (e) providing direct feedback to intervention agents during implementation. Table III presents an intervention script for listening previewing, a strategy targeting reading achievement. Treatment integrity can be assessed by direct observation, videotaping or audiotaping intervention sessions, or teacher-, parent-, or student-completed fidelity checklists.

Consultants seeking to help teachers implement classroom strategies must be prepared to provide

TABLE III
Listening Previewing[a]

Overview

Listening previewing, also termed *passage previewing,* has been shown to improve oral reading performance for a variety of student populations. In this intervention, the teacher reads the assigned passage aloud while students follow along silently prior to independent reading. Higher reading rates have been obtained with bilingual learners and students with learning disabilities, as well as with students in general education. Listening reviewing has also been shown to be superior in improving reading performance to previewing by the traditional silent reading method. Variations include using peers, parents, or tutors as passage previewers.

Purpose

To enhance reading accuracy and fluency by providing an opportunity for students to hear what they will read prior to reading it for themselves.

Materials

1. Stopwatch or watch with second hand (for assessment component)
2. Sheets of paper or copies of students' reading passages for marking oral reading errors (optional)
3. Notebooks or folders with loose-leaf paper for graphing words read correctly per minute (WCPM) rates, one per student (optional)

Observation (Select one or both)

Option 1

1. Administer curriculum-based measurement (CBM) oral reading probes to the students in a reading group or a group of selected students by having each student read a 2-minute timed sample from the previous day's reading passage.
2. Calculate each student's WCPM rate by subtracting the number of errors from the total number of words read and then dividing that figure by two (for number of minutes).
3. Record and graph (or have students graph) WCPM rates.

Option 2

1. Calculate scores on reading skill sheets, quizzes, or end-of-unit tests for all of the students in a reading group or for selected students for 5 to 10 days or for several weeks.

Procedure

1. Divide the reading lesson into passages according to the number of students in the reading group.
2. Tell the students to follow along as you read the first passage aloud.
3. After completing the first listening preview, have the first student read that passage aloud.
4. Repeat this process until you have previewed the entire reading lesson and each student in the reading group has read aloud. If desired, record the number of errors (mispronunciations, substitutions, and omissions) on your copy of the passage or on a separate sheet of paper.
5. Calculate and graph WCPM rates for students in the reading group as often as possible (at least once a week for several weeks) or teach students to chart their own progress on line graphs in their work folders.

Continues

Evaluation (Select one or both)

Option 1

   1. Compare WCPM rates for the target students before and after implementation.

Option 2

   1. Compare scores on reading skill sheets, quizzes, or end-of-unit tests for the target students before and after implementation.

Variations

   1. Have parent volunteers, aides, or peer tutors conduct the listening previews.
   2. Divide the class into pairs, with a higher performing student paired with a lower performing student. Have the higher performing students read first to serve as models for the lower performing students.
   3. Prerecord (or have aides, parent volunteers, or high-performing students prerecord) selected reading passages for independent listening previews at listening stations in the classroom.
   4. To enhance vocabulary and comprehension, have previewers discuss the meaning of important words in the text with the students prior to having the students read.

Notes

   Including a self-monitoring component by having students graph their WCPM performance increases motivation and enhances the effectiveness of this intervention.

Sources

   Rose, T. L. (1984). The effects of two prepractice procedures on oral reading. *Journal of Learning Disabilities, 17,* 544–548.

   Rose, T. L., & Sherry, L. (1984). Relative effects of two previewing procedures on LD adolescents' oral reading performance. *Learning Disability Quarterly, 7,* 39–44.

   Rousseau, J. K., Tam, B. K., & Ramnarain, R. (1993). Increasing the reading proficiency of language-minority students with speech and language impairments. *Education and Treatment of Children, 16,* 254–271.

   [a]Adapted from *Effective School Interventions: Strategies for Enhancing Academic Achievement and Social Competence* (pp. 168–170) by N. Rathvon, 1999, New York: Guilford Press. Copyright 1999 by Guilford Press. Adapted with permission.

support with many aspects of the intervention process. Although teachers must implement the strategy, studies suggest that treatment integrity and, ultimately, the success of the intervention are related to the degree to which classroom-based assistance is provided to teachers during implementation. For example, IAPs in which a designated case manager works collaboratively with the referring teacher throughout the implementation process are more likely to report measurable gains in student performance than are IAPs lacking such support systems.

4.3. Evaluating Academic Interventions

Systematically evaluating performance change not only provides information that is useful in monitoring and increasing intervention effectiveness but also contributes to teachers' maintenance of interventions by demonstrating that positive change is occurring.

Although researchers have developed several measures for assessing teachers' perceptions of changes in students' academic performance, it is important to assess actual student outcomes and not merely teachers' or parents' perceptions of improvement. Similarly, researchers have often evaluated the effects of academic interventions in terms of task completion rates without regard for accuracy or the absolute level of achievement. Ultimately, the effectiveness of academic interventions should be evaluated in terms of meaningful changes in students' academic achievement relative to grade-level expectations.

4.3.1. Curriculum-Based Assessment

In recent years, intervention-oriented researchers have developed alternative assessment methodologies to traditional norm-referenced tests with the goal of identifying students in need of supplementary academic services and documenting the effectiveness of school-based interventions. One of these methods, curriculum-based assessment (CBA), refers to a set of procedures

that link assessment directly to instruction and evaluate progress using measures taken from the students' own curricula. Among the many different CBA models, the most fully developed is curriculum-based measurement (CBM), which has become the standard for assessing changes in student performance subsequent to interventions, especially in reading. Developed by Deno, Mirkin, and colleagues at the University of Minnesota Institute for Learning Disabilities, CBM is a generic measurement system that uses brief, fluency-based measures of basic skills in reading, mathematics, spelling, and written expression. CBM is ideally suited to monitoring the progress of students receiving academic interventions because measures are brief (1–3 min), can be administered frequently, and are based on students' own instructional materials. Procedures for conducting CBMs in reading, mathematics, spelling, and written expression can be found in Rathvon's 1999 book.

5. ACADEMIC INTERVENTION TARGETS

Academic interventions can be categorized according to targets as follows: (a) interventions designed to enhance academic productivity, including classwork, independent seatwork, and homework; (b) interventions targeting achievement in specific academic subjects; and (c) interventions targeting what DiPerna and Elliott termed academic enablers, that is, non-academic skills, behaviors, and attitudes that contribute to academic competence. Table I describes empirically validated academic interventions from each of the three categories. The categorization is necessarily somewhat arbitrary because all of the interventions include procedures that facilitate productivity and academic enabling behaviors; however, the interventions in the academic enablers category include the largest number of behavioral and attitudinal components.

5.1. Interventions Targeting Academic Productivity

5.1.1. Self-Management

Self-management techniques involve teaching students to engage in some form of behavior, such as self-observation or self-recording, in an effort to alter a target behavior. Self-management interventions fall into one of two categories: (a) contingency-based strategies with self-reinforcement for the performance of

specified tasks or (b) cognitively based strategies that use self-instruction to address academic deficits. Self-management interventions are especially appropriate for targeting academic problems because they not only enhance students' sense of responsibility for their own behavior but also increase the likelihood that students will be able to generalize their new competencies to other situations. Many academic interventions, including most of those listed in Table I, include at least one self-management component.

5.1.2. Classwide Peer Tutoring

Increased academic responding is associated with higher levels of on-task behavior and achievement. In classwide peer tutoring, peers supervise academic responding so that every student can engage in direct skill practice during instructional periods, leaving teachers free to supervise the tutoring process. Moreover, because peer tutors are provided with the correct answers for tutoring tasks, the strategy permits immediate error correction. Of the several variations of this strategy, the best known is the Classwide Peer Tutoring (CWPT) program developed by Greenwood and colleagues at the University of Kansas to improve the achievement of entire classrooms of low-socioeconomic status urban students. CWPT has been successful in improving academic skills and productivity in a variety of domains, including oral reading, spelling, and mathematics computation, and with both regular education and special needs students.

5.2. Interventions Targeting Academic Achievement

This section discusses some of the best-known and most widely validated interventions in three academic areas: reading, mathematics, and written language.

5.2.1. Interventions to Improve Reading Performance

Reading problems are the most frequent cause of referrals to school psychologists and IAPs. Three sets of skills are required for proficient reading: (a) decoding (i.e., the process leading to word recognition), (b) comprehension (i.e., the ability to derive meaning from text), and (c) fluency (i.e., the ability to read quickly and accurately). Although reading interventions can be categorized according to their primary subskill target, interventions focusing on one subskill

have the potential to improve other competencies due to the interrelated nature of the reading process. The two interventions described in this section primarily target fluency and comprehension.

5.2.1.1. Listening Previewing Listening previewing, one of the best-known stand-alone reading interventions, promotes fluency by providing students with an effective reading model prior to the students reading aloud. Combining listening previewing with discussion of key words in the selection to be read is associated with enhanced outcomes, especially in terms of comprehension. Listening previewing ranks high in both efficacy and usability and has been implemented successfully with regular education students, bilingual students, and students with learning and behavior disorders. Additional information about listening previewing is provided in Table III.

5.2.1.2. Collaborative Strategic Reading Collaborative strategic reading (CSR) combines instruction in comprehension strategies and study skills with collaborative peer practice. Students learn four strategies through direct instruction and teacher modeling: (a) Preview (i.e., previewing and predicting), (b) Click and Clunk (i.e., monitoring for understanding and vocabulary knowledge), (c) Get the Gist (i.e., understanding the main idea), and (d) Wrap-Up (i.e., self-questioning for understanding). After students have mastered the strategies, they implement them within cooperative groups. CSR has been successful in improving reading proficiency in regular education, multilevel, inclusive, and special education settings. Originally designed for use with expository text in content area textbooks, it can also be applied to narrative material.

5.2.2. Interventions to Improve Mathematics Performance

The 2003 NAEP report on mathematics revealed serious deficiencies in math achievement in the general student population. Although the percentage of fourth graders performing at or above the proficient level (29%) was higher in 2003 than in all previous assessment years since 1990, sizable numbers of students failed to reach even the basic level of proficiency (23%). The situation was even more dismal for eighth graders, with only 23% performing well enough to be classified as proficient. Although students with difficulties in learning mathematics constitute a very heterogeneous group, they generally exhibit deficits in one or more of three areas:

(a) computational skills, including the basic operations of addition, subtraction, multiplication, and division; (b) computational fluency, that is, speed and automaticity with math facts; and (c) mathematics applications, including areas such as money, measurement, time, and word problems. Targets of mathematics interventions can be characterized as (a) foundational arithmetic skills (e.g., number knowledge, basic understanding of mathematical operations), (b) acquisition and automatization of basic computational skills, and (c) problem-solving skills. Recent meta-analyses of mathematics interventions with low-achieving students and/or students with disabilities have reported the greatest efficacy for interventions targeting basic skills and the least efficacy for interventions focusing on higher order mathematics skills. Promising intervention components include frequent feedback to teachers and parents regarding student performance, explicit instruction in math concepts and procedures, and peer-assisted learning. The two interventions described in this section use highly structured, multicomponent approaches to enhance basic skills acquisition and fluency.

5.2.2.1. Cover–Copy–Compare The cover–copy–compare (CCC) strategy, originally developed by Skinner and colleagues at Mississippi State University, is a self-management intervention that can be used to enhance accuracy and fluency in a variety of academic subjects. Students look at an academic stimulus (e.g., a multiplication problem for CCC mathematics), cover it, copy it, and evaluate their response by comparing it to the original stimulus. CCC combines several empirically based intervention components, including self-instruction, increased opportunities to respond to academic material, and immediate corrective feedback.

5.2.2.2. Reciprocal Peer Tutoring Reciprocal peer tutoring (RPT) in mathematics, developed by Fantuzzo and associates at the University of Pennsylvania, combines self-management techniques and group contingencies within a peer tutoring format. Although both CWPT and RPT involve peer-mediated instruction, RPT includes self-management and subgroup contingencies, with teams of students selecting and working to obtain their own rewards. RPT has been demonstrated to enhance not only math performance but also students' perceptions of their own scholastic competence. Including a home-based reinforcement component enhances positive outcomes.

5.2.3. Interventions to Improve Written Language Performance

Writing is a crucial skill for school success because it is a fundamental way in which to communicate ideas and demonstrate knowledge in the content areas. Unfortunately, writing problems not only are characteristic of most students with learning disabilities but also are prevalent in the general student population. According to the 2002 NAEP report on writing, only 26% of 4th graders scored at the proficient level, with the percentage dropping to 22% for 12th graders. The pervasiveness of writing problems suggests that poor writing achievement is related less to internal student disabilities than to inadequate writing instruction. During the past decade or so, research on the cognitive processes underlying writing has led to a shift in writing instruction from an emphasis on product (e.g., grammar, mechanics, content) to an emphasis on the processes used to generate written productions (e.g., brainstorming, writing multiple drafts, developing a sense of audience, incorporating feedback from others). As a result, writing interventions increasingly focus on student performance of various aspects of the writing process, including planning, sentence generation, and revising. The interventions in this section target several writing process components, including fluency and compositional elements such as planning and editing.

5.2.3.1. Self-Regulated Strategy Development in Writing Self-regulated strategy development (SRSD) is an instructional approach that combines explicit teaching and modeling of compositional strategies with a set of self-regulation procedures. In a series of studies, Graham, Harris, and colleagues at the University of Maryland have documented that SRSD improves the quantity and quality of narrative and expository writing for students with writing disabilities as well as for normally developing writers. In addition to enhancing writing performance, SRSD is associated with improved feelings of self-efficacy and more positive attitudes toward writing for students. Most applications of SRSD include paired and/or small-group learning activities at specific steps in the training process to provide additional opportunities for feedback and collaboration without direct teacher supervision.

5.2.3.2. Self-Recording of Writing Productivity This intervention uses CBM-type methodology in the context of self-monitoring procedures to increase writing fluency. Self-recording word counts during free writing periods provides students with opportunities for positively evaluating their own writing fluency and yields useful progress monitoring data by showing teachers when changes in instruction are followed by increases in writing performance rates. Setting writing production goals and self-recording word counts produce increases in the number of words written and improvements in free writing expressiveness without a deterioration in writing mechanics, compared with untimed writing periods. Some variations include individual and group reinforcements for achieving preset goals.

5.3. Interventions to Enhance Academic Enablers

According to DiPerna and Elliott, academic success requires more than skill in performing assigned tasks. That is, although classroom instruction focuses on the acquisition of concepts, knowledge, and skills in academic subjects, students must become active participants in their educational experiences to benefit from that instruction. The two interventions in this category have been widely used as stand-alone strategies or in combination with other intervention components to facilitate achievement and productivity across a range of academic skill domains.

5.3.1. Public Posting

Public posting involves displaying some kind of classroom record (e.g., a chart) that documents student achievement or productivity. Whereas traditional public posting strategies often record incidents of negative behavior (e.g., the names of disruptive or unproductive students), public posting as an academic intervention is a positive strategy that displays student progress in achieving specified academic goals. Originally designed to encourage improvement in individual student performance, small-group and classwide versions that target group achievement and capitalize on positive peer influence have also been developed. Adding individual or group contingencies can enhance outcomes but does not appear to be critical to intervention effectiveness.

5.3.2. School–Home Notes

School–home notes encourage parental involvement in children's classroom performance, permit a broader range of reinforcers than are generally available to teachers, and have demonstrated efficacy for both

academic problems and behavior problems. School–home communications can be arrayed along a continuum of parental involvement, ranging from notes that merely provide information to notes that ask parents to deliver predetermined consequences contingent on the reported student performance. Although strategies that include home consequences can have powerful effects on student performance, establishing and maintaining an effective school–home communication system can be difficult for even a single student, much less for groups or entire classrooms of unproductive students. Not surprisingly, the majority of published school–home note interventions have targeted individual students or small groups of students, usually in special education settings.

6. SUMMARY AND FUTURE DIRECTIONS

Given the growing diversity and needs of the student population, interest in the development and empirical validation of academic interventions is likely to increase. Although the knowledge base of effective school-based interventions has increased dramatically during the past 15 years, determining which intervention components and which intervention parameters (e.g., intensity, duration, delivery system) are maximally effective for which types of students will be a continuing challenge for researchers and practitioners alike. Additional studies are especially needed to identify strategies that are high in both efficacy and usability for second-language learners and secondary school students with special needs.

See Also the Following Articles

Academic Failure, Prevention of ■ Advertising and Culture ■ Behavioral Assessment in Schools ■ Educational Achievement and Culture ■ Educational and Child Assessment ■ Effective Classroom Instruction ■ Learning Styles and Approaches to Studying ■ Mathematics, Teaching of ■ Reading Interventions ■ Reading, Teaching of ■ Teaching Effectiveness ■ Transfer of Learning ■ Writing, Teaching of

Further Reading

Daly, E. J., & McCurdy, M. (Eds.). (2002). Developments in academic assessment and intervention [Special issue]. *School Psychology Review*, *31*(4).

DiPerna, J. C., & Elliott, S. N. (Eds.). (2002). Promoting academic enablers to improve student performance: Considerations for research and practice [Miniseries]. *School Psychology Review*, *31*(3).

Elliott, S. N., Busse, R. T., & Shapiro, E. S. (1999). Intervention techniques for academic performance problems. In C. R. Reynolds, & T. B. Gutkin (Eds.), *Handbook of school psychology* (3rd ed., pp. 664–685). New York: John Wiley.

Gutkin, T. B. (Ed.). (2002). Evidence-based interventions in school psychology: The state of the art and future directions [special issue]. *School Psychology Quarterly*, *17*(4).

Rathvon, N. (1996). *The unmotivated child: Strategies for helping your underachiever become a successful student*. New York: Simon and Schuster.

Rathvon, N. (1999). *Effective school interventions: Strategies for enhancing academic achievement and social competence*. New York: Guilford Press.

Shinn, M. R., Walker, H. M., & Stoner, G. (Eds.). (2002). *Interventions for academic and behavior problems II*: *Preventive and remedial approaches*. Bethesda, MD: National Association of School Psychologists.

Swanson, H. L. (2000). What instruction works for students with learning disabilities? Summarizing the results from a meta-analysis of intervention studies. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 1–30). Mahwah, NJ: Lawrence Erlbaum.

Upah, K. R. F., & Tilly, W. D. (2002). Best practices in designing, implementing, and evaluating quality interventions. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 483–502). Bethesda, MD: National Association of School Psychologists.

Accidents in Transportation

Sidney Dekker

Linköping Institute of Technology, Linköping, Sweden

GLOSSARY

accident Generally unforeseen occurrence with damaging consequences for life, limb, or property.

automation Replacement of human work by machines or computers in many transportation modes; it often creates new human work in form of new knowledge and skill demands and has been linked to the emergence of new kinds of accidents.

chain of events Model of accident causation that captures the multiple subsequent failures necessary to cause an accident (i.e., domino effect); it has recently been challenged by systemic and control models.

regulation Activity of defining rules to govern (transportation) practice and the enforcement of those rules; this often requires large organizations that are nearly always government sponsored and run.

systemic and control models Models that see accidents as a result of loss of control or as a gradual erosion of safety constraints on system design and operation.

Transportation accidents carry great personal, organizational, societal, and economic costs. Applied psychology can derive research inspiration from, and can contribute much to the understanding and eventual prevention of, accidents in all kinds of transportation modes. The applied psychological questions surrounding accidents can vary widely, from issues of perception in an individual driver, to trade-offs made in complex organizations under pressures of scarcity and competition, to biases and heuristics in our understanding of culpability, cause, and control. Further progress on safety in transportation systems hinges in part on contributions from applied psychology.

1. WHY IS IT IMPORTANT TO UNDERSTAND ACCIDENTS?

Transportation accidents carry tremendous social, environmental, and economic costs. Traffic accidents are among the leading causes of death for people under 34 years of age. In 2002 alone, 42,815 people died in highway accidents in the United States; that equates to more than 117 fatalities per day or 1.51 fatalities per 100 million vehicle miles traveled. The number of injuries hovers around 3 million per year. The economic costs of these accidents are enormous—more than $230 billion per year for the United States alone. Survivors of transportation accidents often require treatment and rehabilitation and also suffer psychological consequences that may carry over into work and social settings.

The nature and frequency of accidents differ widely among the various transportation modes. For example, the year 1999 witnessed 51 major accidents in commercial aviation and 2768 train-related accidents that were responsible for 932 casualties. At the same time, safety of the various transportation modes varies with geographical location. For example, in terms of casualties, passenger shipping may be less safe than driving in parts of Southeast Asia, and flying may be more dangerous in certain parts of Africa. Indeed, it would seem that the phrase "richer is safer" is true except when considering road traffic accidents in heavily populated, affluent parts of world. All the same, a transportation accident is one of the most common risks to which any member of society is exposed.

2. APPLIED PSYCHOLOGY AND ACCIDENTS

Transportation accidents are interesting from an applied psychology perspective because they are, most basically, about human behavior in applied settings. In fact, accidents are typically about human cognition as much as they are about technical problems and issues. The applied psychological questions surrounding accidents can vary widely, from issues of sensation and perception in an individual driver to shortcuts made by organizational members under pressures of scarcity and competition. Applied psychology as a field of scientific inquiry can use its insights for, and derive much empirical inspiration from, transportation accidents. Indeed, applied psychology in the broadest, most inclusive sense is relevant to transportation accidents.

In studying whether a car driver could have seen a pedestrian, applied psychology may be called on to assess the effect of atmospheric modulation (e.g., rain, mist, dust) on stimulus perceptibility (e.g., the pedestrian). Models exist to calculate so-called atmospheric modulation transfer that can help applied psychologists to reconstruct the existing visibility during a particular occurrence. Additional questions may arise when substance abuse is suspected on the part of the driver; in such a case, applied psychology relies more on models of psychopathology. Much research leverage exists in gauging the (sub)cultural acceptability of consuming alcohol before engaging in transportation activities. Other questions with relevance to applied psychology arise when notable mismatches occur in people's risk perceptions. For example, people might

not want to use the railroads due to one highly publicized recent fatal accident. If they prefer to drive their cars, they actually expose themselves to the risk of an accident (perhaps even a fatal accident) several orders of magnitude larger than that which would exist while riding the train. Such individual risk assessments are deeply confounded, mixing personal histories, notions of control and destiny, and various kinds of biases and rationalizations (e.g., the availability heuristic) into a rich trove for applied psychology.

Perhaps even more complex are questions surrounding the trade-offs made by professionals who work in transportation systems, leading to issues such as "safety cultures" in transportation systems. Does the commander of the aircraft continue with the approach in bad weather or not? Does the master of the ship sail in the storm or not? Such trade-offs are more than the risk assessments of individual decision makers; they must also be understood as expressions of the preferences and priorities of entire sociotechnical systems that operate in environments of scarcity (e.g., only so many passengers, only so much money to be had) and competition (e.g., somebody else is always ready to take over one's routes). In efforts to investigate these kinds of trade-offs and other applied problems, psychology has seen a methodological shift accelerate over the past decade or so, with an increasing emphasis on fieldwork and the study of applied settings. It is believed that to understand decision making in actual (transportation) work, where real decisions have real outcomes for real people, psychological researchers must get out of the laboratory and investigate applied settings directly. Issues of confounding factors are dealt with through intensive analysis and interpretation of research results, leading to findings (e.g., on naturalistic decision making) that are both internally valid and exportable to other applied settings.

There is also a role for applied psychology in understanding issues of culpability and control. To what extent, and why, do we judge participants in transportation (e.g., drivers, pilots) to be culpable for the accidents that they "cause"? Such questions are deeply complex and touch on much of what people hope and believe about the world in which they live. The questions are interconnected with yet another set of biases, including the "hindsight bias" that describes how knowledge of outcome profoundly alters one's perception of the behavior and intentions that led up to that outcome. The hindsight bias also allows one to convert a complex tangled history leading up to an accident into a simple series of binary decisions, where

participants had clear choices to do the right thing or the wrong thing. These questions also inevitably coincide with assumptions about causation, whether justified or not. These tend to be quite problematic and affect applied psychological reasoning around accidents and human intention and behavior.

3. WHY DO ACCIDENTS HAPPEN?

One of the more compelling questions for those involved in transportation is why accidents happen. Indeed, a pressing issue after an accident is often to resolve the question, "What was the cause?" There is, of course, no one or unequivocal answer, and not just because the various modes of transportation may differ widely in their respective etiologies of failure. Models of accident causation develop continually, reflecting not only new insights or access to accident data but also the general scientific spirit of the times. In transportation, the so-called "chain of events" model is popular. In this model, one failure somewhere in the system can be seen to lead to (or trigger) the next, and so on, until this cascade of individually insignificant faults pushes the entire system over the edge of breakdown. For example, a piece of debris on the runway, left there by a preceding aircraft, manages to puncture the tire of a subsequent aircraft. Tire fragments slam into the wing, puncturing the fuel tank and triggering a fire that, not much later, brings down the entire aircraft. The chain of events model is also credited with distributing or refocusing the search for accident causes away from frontline operators (i.e., away from "operator error"), instead identifying higher up supervisory and organizational shortcomings that may have contributed.

However, recent insight into the sociotechnical nature of large transportation accidents, such as disasters involving the Space Shuttles *Challenger* (in 1986) and *Columbia* (in 2003), has exposed the limits of the chain of events model. Most critically, the model presupposes (or even requires) failures so as to cause a failure. This contradicts characterizations of organizational and operational practice preceding accidents as normal everyday routine. Accidents can still happen even if everybody follows the rules and there are no "failures" or "shortcomings" as seen from the perspective of those running and regulating the transportation system. As Perrow suggested in 1984, accidents in these systems may be quite "normal" given their structural properties of interactive complexity and tight coupling. Nothing extraordinary (i.e., no obvious breakdowns or failures) as seen

from inside the system is necessary to produce an accident. For example, aviation is a tightly coupled system, as are railroads. This means that there is little slack to recover if things do start to go wrong. The workings of a railroad system, however, are generally more linear and transparent than those of aviation, with the former allowing for better insight into how to manage the system.

These insights have put pressure on the old label "human error" as an explanation for accidents. Indeed, human error today can no longer be legitimately seen as the cause of accidents; rather, it is seen as an effect, as a symptom, or as a sign of trouble deeper inside the system. Human error is no longer an explanation; instead, it demands one. Human error, to the extent that it exists as a separable category of human performance (which many researchers doubt or deny), is systematically connected to features of people's tools and tasks. This is consistent with the ecological commitment in much of applied psychology, where an analysis of agent–environment mutuality, or person–context couplings, is critical to producing useful insights into the success and failure of transportation (and other) systems.

The latest accident models deemphasize "cause" altogether, noting that it is a deeply Newtonian concept that might not carry over well into an understanding of why complex transportation systems fail, with the *Columbia* accident being a case in point. These models see transportation accidents rather as control problems; as the gradual erosion and eventual loss of control over a safety-critical process, where safety constraints on design or operation are violated. Such models can be reconciled better with increasing evidence that large transportation accidents are nearly invariably preceded by some kind of slow but gradual drift (e.g., away from procedures, away from design specifications) that is hard to notice or characterize as deviant when seen from the inside, that is, from the perspective of participants themselves. The consistent, if slight, speeding that most male drivers exhibit (e.g., nearly always 5 miles per hour over the speed limit) is another example of this. The deviance is normalized; the nonroutine becomes routine, and there is consistent noncompliance. Production pressures (e.g., pressure to be on time, pressure to maximize capacity use) form the major engine behind such drift toward failure in nearly all transportation systems because virtually no transportation system is immune to the pressures of competition and resource scarcity. Recent research debates are not necessarily about the existence or reality of such pressures; rather, they discuss to what extent these pressures consciously affect people's

trade-offs and to what extent people deliberately gamble (and lose if they have an accident). Much research points to the prerational insidious nature of production pressures on people's trade-offs rather than to conscious immoral or risky calculation, although male drivers under 25 years of age might be an exception.

Acknowledging that such sociotechnical complexity underlies most transportation accidents means that safety and risk are really social constructs rather than objective engineering measurements. Risk is constructed at the intersection of social forces and technical knowledge (or the lack thereof). Slogans such as "safety comes first" are mere posturing. They are disconnected from reality where risk is negotiated as subjective social activity and where it will be different in different cultures, different among organizations within a single culture, and even different among subcultures within one organization.

4. REGULATION OF TRANSPORTATION SYSTEMS AND INCIDENT REPORTING

Most, if not all, transportation systems are regulated to some extent. Regulation means that rules are created to govern practice and that the rules are enforced to generate compliance. This often requires large organizations—nearly always public or government agencies—tasked with issuing rules, overseeing practice, and certifying and checking systems, operators, and organizations. The extent of regulation of various transportation modes varies greatly from country to country. In some countries, regulators are criticized for having too close a relationship with the transportation industries they are supposed to regulate and for having a dual and supposedly incompatible mandate (i.e., promoting the use and growth of the transportation system as well as overseeing and monitoring its compliance).

The success of safety regulation depends on the safety level that the transportation system has already achieved. In relatively unsafe transportation modes (e.g., private flying), regulation can pay great and rather immediate dividends. It can help to standardize practice; it can issue, broadcast, and enforce rules and remind operators of them; and it can help to build a corpus of cases on routes to accidents that can be shared in the community. There is often leverage in changing designs (e.g., operator interfaces) to make them more error

resistant and error tolerant. In safer systems (e.g., charter flights), such error-resistant designs have typically evolved further. Accidents are often preceded by so-called "dress rehearsals"—sequences of events similar to real accidents but without fatal outcomes. Learning from those dress rehearsals is encouraged, particularly through incident-reporting systems. However, in ultra-safe transportation systems (e.g., European railroads), overregulation becomes a problem. More rules and more monitoring are no longer accompanied by safety gains; they serve only to increase system complexity and potentially decrease transparency as well as rule compliance. Incident-reporting systems in ultra-safe transportation systems may be of limited value. The typical accident there emerges from routine, everyday banal factors that combine and align in ways that are hard to foresee. Incident reports would neither notice the relevance of those factors nor be able to subsequently project their interplay. This, in fact, is a problem allied with all incident-reporting systems. Reporting is not the same as analysis, which in turn is not the same as learning from potential failure. Just having an incident reporting system in place guarantees little in terms of progress on safety.

5. INTERVENTIONS, NEW TECHNOLOGY, AND FUTURE DEVELOPMENTS

Prevention and improvement strategies vary considerably from one transportation mode to another because many of the factors that affect accident likelihood are mode specific. In general, prevention strategies seek either to minimize the likelihood of a particular kind of harmful event (e.g., a crash) or to minimize its impact (in terms of property damage, injury, or environmental pollution). New technology often plays a dominant role here. Automatic seat belt systems, side impact protection systems, collision detection and avoidance systems, anti-lock breaking systems, double-hull vessels, automated cocoons to keep an aircraft within its prescribed envelope—all of these systems intervene, prevent, and protect in one sense or another. Even though enormous progress is made through such interventions, technology is sometimes a mixed blessing. Airbags may protect mostly drivers who do not wear seat belts (forcing a large majority of other people, who do wear seat belts, to pay a premium on their cars), and automation in virtually all transportation

modes has been associated with the emergence of new human–machine coordination problems (e.g., mode errors, "automation surprises," people getting lost in display page architectures). Indeed, automation has been associated with the emergence of new types of accidents in virtually all transportation modes, where it has introduced new capabilities and new complexities. One typical signature of accidents in automated transportation systems is the "going sour" scenario, where a small trigger event (i.e., an unusual occurrence or a small failure in a system somewhere), itself innocuous, leads to a series of misassessments and miscommunications between humans and computers. The system is managed into greater hazard due to coordination breakdowns among the various players (both humans and machines). The end result is often an automation surprise, where humans discover that what they thought they had instructed the automation to do was not quite what the automation was doing. More research, much of it in applied psychology, will be necessary to use new technology and other interventions to their full potential in helping to reduce transportation accidents.

See Also the Following Articles

Aviation ■ Driving Safety ■ Models for Transportation ■ Traffic Safety Assessment ■ Transportation Systems, Overview

Further Reading

Amalberti, R. (2002). The paradoxes of almost totally safe transportation systems. *Safety Science*, *37*, 109–126.

Billings, C. E. (1996). *Aviation automation: The search for a human centered approach.* Mahwah, NJ: Lawrence Erlbaum.

Dekker, S. W. A. (2002). *The field guide to human error investigations.* Bedford, UK: Cranfield University Press.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies.* New York: Basic Books.

Vaughan, D. (1996). *The* Challenger *launch decision: Risky technology, culture, and deviance at NASA.* Chicago: University of Chicago Press.

Acculturation

J. W. Berry

Queen's University, Kingston, Ontario, Canada

GLOSSARY

acculturation A process of cultural and psychological change resulting from contact between cultural groups and their individual members.

acculturation strategies Variations in the way acculturating groups and individuals attempt to manage the process of acculturation (see Assimilation, Integration, Marginalization, Separation).

acculturative stress A negative psychological reaction to the experiences of acculturation, often characterized by anxiety, depression and a variety of psychosomatic problems.

adaptation The long-term outcome of acculturation, including psychological adaptation (feelings of personal well-being and self-esteem) (and sociocultural adaptation (competence in dealing with life in the new society).

assimilation The acculturation strategy in which people do not wish to maintain their heritage culture, and seek to participate in the larger society.

cultural distance The degree of cultural dissimilarity between two groups, measured by ethnographic indicators, or by an individual's perception of such difference.

cultural identity How individuals think and feel about themselves in relation to their two cultural groups in contact.

integration The acculturation strategy in which people wish to maintain their heritage culture, and seek to participate in the larger society.

larger society An evolving social and political framework representing a civic consensus about how to live together in a culturally plural society.

marginalization The acculturation strategy in which people do not maintain their heritage culture, and also do not participate in the larger society.

multicultural ideology A positive orientation of individuals to cultural diversity in plural societies involving the acceptance of ethnocultural groups and their participation in the larger society.

plural society A society in which a number of ethnocultural groups live together within a shared civic political and economic framework.

psychological acculturation A process in which psychological features of a person change as a result of their contact with other cultural groups.

separation The acculturation strategy in which people maintain their heritage culture, and seek to avoid participation in the larger society.

Acculturation is the process of cultural and psychological change that takes place as a result of contact between cultural groups and their individual members. Such contact and change occur during colonization, military invasion, migration, and sojourning (e.g., tourism, international study, overseas posting). It continues after initial contact in culturally plural societies, where ethnocultural communities maintain features of their heritage cultures. Adaptation to living in culture contact

settings takes place over time. Occasionally it is stressful, but often it results in some form of accommodation. Acculturation and adaptation are now reasonably well understood, permitting the development of policies and programs to promote successful outcomes for all parties.

1. ACCULTURATION CONCEPT

The initial interest in acculturation grew out of a concern for the effects of European domination of colonial and indigenous peoples. Later, it focused on how immigrants (both voluntary and involuntary) changed following their entry and settlement into receiving societies. More recently, much of the work has been involved with how ethnocultural groups relate to each other, and to change, as a result of their attempts to live together in culturally plural societies. Nowadays, all three foci are important as globalization results in ever larger trading and political relations. Indigenous national populations experience neocolonization; new waves of immigrants, sojourners, and refugees flow from these economic and political changes; and large ethnocultural populations become established in most countries.

Early views about the nature of acculturation are a useful foundation for contemporary discussion. Two formulations in particular have been widely quoted. The first, from Redfield and colleagues in a 1936 article, is as follows:

> Acculturation comprehends those phenomena which result when groups of individuals having different cultures come into continuous first-hand contact, with subsequent changes in the original culture patterns of either or both groups. ... Under this definition, acculturation is to be distinguished from culture change, of which it is but one aspect, and assimilation, which is at times a phase of acculturation.

In another formulation, the Social Science Research Council in 1954 defined acculturation as

> culture change that is initiated by the conjunction of two or more autonomous cultural systems. Acculturative change may be the consequence of direct cultural transmission; it may be derived from non-cultural causes, such as ecological or demographic modification induced by an impinging culture; it may be delayed, as with internal adjustments following upon the acceptance of alien traits or patterns; or it may be a reactive adaptation of traditional modes of life.

In the first formulation, acculturation is seen as one aspect of the broader concept of culture change (that

which results from intercultural contact), is considered to generate change in "either or both groups," and is distinguished from assimilation (which may be "at times a phase"). These are important distinctions for psychological work and are pursued later in this article. In the second definition, a few extra features are added, including change that is indirect (not cultural but rather "ecological"), is delayed (internal adjustments, presumably of both a cultural and a psychological character, take time), and can be "reactive" (i.e., rejecting the cultural influence and changing toward a more "traditional" way of life rather than inevitably toward greater similarity with the dominant culture).

In 1967, Graves introduced the concept of psychological acculturation, which refers to changes in an individual who is a participant in a culture contact situation, being influenced both by the external culture and by the changing culture of which the individual is a member. There are two reasons for keeping these two levels distinct. The first is that in cross-cultural psychology, we view individual human behavior as interacting with the cultural context within which it occurs; hence, separate conceptions and measurements are required at the two levels. The second is that not every individual enters into, and participates in, a culture in the same way, nor does every individual change in the same way; there are vast individual differences in psychological acculturation, even among individuals who live in the same acculturative arena.

A framework that outlines and links cultural and psychological acculturation and identifies the two (or more) groups in contact is presented in Fig. 1. This framework serves as a map of those phenomena that the author believes need to be conceptualized and measured during acculturation research. At the cultural level (on the left of the figure), we need to understand key features of the two original cultural groups (A and B) prior to their major contact, the nature of their contact relationships, and the resulting dynamic cultural changes in both groups, and in the emergent ethnocultural groups, during the process of acculturation. The gathering of this information requires extensive ethnographic, community-level work. These changes can be minor or substantial and can range from being easily accomplished to being a source of major cultural disruption. At the individual level (on the right of the figure), we need to consider the psychological changes that individuals in all groups undergo and their eventual adaptation to their new situations. Identifying these changes requires sampling a population and studying individuals who are variably
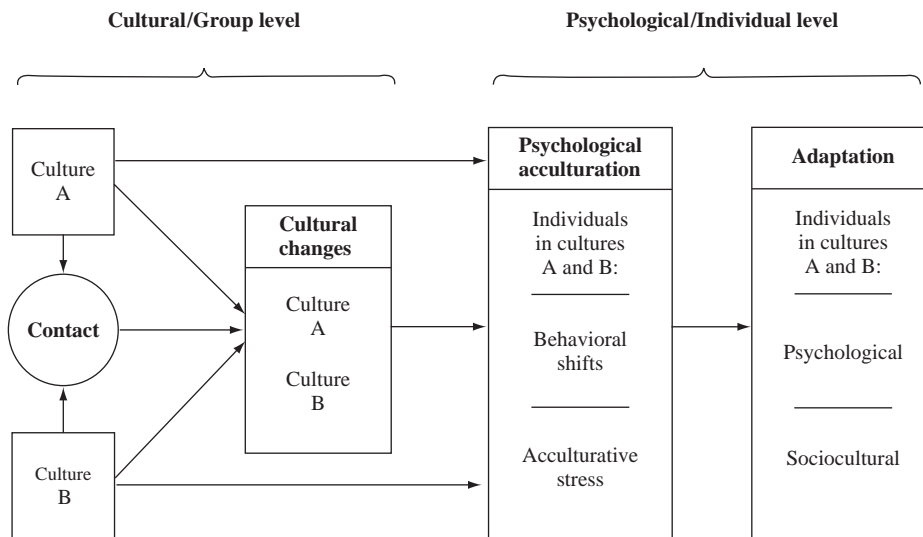
FIGURE 1   A general framework for understanding acculturation.

involved in the process of acculturation. These changes can be a set of rather easily accomplished behavioral shifts (e.g., in ways of speaking, dressing, or eating; in one's cultural identity), or they can be more problematic, producing acculturative stress as manifested by uncertainty, anxiety, and depression. Adaptations can be primarily internal or psychological (e.g., sense of well-being or self-esteem) or sociocultural, linking the individual to others in the new society as manifested, for example, in competence in the activities of daily intercultural living.

2. ACCULTURATION CONTEXTS

As for all cross-cultural psychology, it is imperative that we base our work on acculturation by examining its cultural contexts. We need to understand, in ethnographic terms, both cultures that are in contact if we are to understand the individuals that are in contact.

In Fig. 1, we saw that there are five aspects of cultural contexts: the two original cultures (A and B), the two changing ethnocultural groups (A and B), and the nature of their contact and interactions.

Taking the immigration process as an example, we may refer to the society of origin (A), the society of settlement (B), and their respective changing cultural features following contact (A1 and B1). A complete understanding of acculturation would need to start with a fairly comprehensive examination of the societal contexts. In the society of origin, the cultural characteristics

that accompany individuals into the acculturation process need description, in part to understand (literally) where the person is coming from and in part to establish cultural features for comparison with the society of settlement as a basis for estimating an important factor to be discussed later: cultural distance. The combination of political, economic, and demographic conditions being faced by individuals in their society of origin also needs to be studied as a basis for understanding the degree of voluntariness in the migration motivation of acculturating individuals. Arguments by Richmond in 1993 suggested that migrants can be arranged on a continuum between reactive and proactive, with the former being motivated by factors that are constraining or exclusionary (and generally negative in character), and the latter being motivated by factors that are facilitating or enabling (and generally positive in character). These contrasting factors were also referred to as push/pull factors in the earlier literature on migration motivation.

In the society of settlement, a number of factors have importance. First, there are the general orientations that a society and its citizens have toward immigration and pluralism. Some societies have been built by immigration over the centuries, and this process may be a continuing one, guided by a deliberate immigration policy. The important issue to understand for the process of acculturation is both the historical and attitudinal situations faced by migrants in the society of settlement. Some societies are accepting of cultural pluralism resulting from immigration, taking steps to

support the continuation of cultural diversity as a shared communal resource. This position represents a positive multicultural ideology and corresponds to the integration strategy. Other societies seek to eliminate diversity through policies and programs of assimilation, whereas still others attempt to segregate or marginalize diverse populations in their societies. Murphy argued in 1965 that societies that are supportive of cultural pluralism (i.e., with a positive multicultural ideology) provide a more positive settlement context for two reasons. First, they are less likely to enforce cultural change (assimilation) or exclusion (segregation and marginalization) on immigrants. Second, they are more likely to provide social support both from the institutions of the larger society (e.g., culturally sensitive health care, multicultural curricula in schools) and from the continuing and evolving ethnocultural communities that usually make up pluralistic societies. However, even where pluralism is accepted, there are well-known variations in the relative acceptance of specific cultural, "racial," and religious groups. Those groups that are less well accepted experience hostility, rejection, and discrimination, one factor that is predictive of poor long-term adaptation.

3. ACCULTURATION STRATEGIES

Not all groups and individuals undergo acculturation in the same way; there are large variations in how people seek to engage the process. These variations have been termed acculturation strategies. Which strategies are used depends on a variety of antecedent factors (both cultural and psychological), and there are variable consequences (again both cultural and psychological) of these different strategies. These strategies consist of two (usually related) components: attitudes and behaviors (i.e., the preferences and actual outcomes) that are exhibited in day-to-day intercultural encounters.

The centrality of the concept of acculturation strategies can be illustrated by reference to each of the components included in Fig. 1. At the cultural level, the two groups in contact (whether dominant or nondominant) usually have some notion about what they are attempting to do (e.g., colonial policies, motivations for migration), or of what is being done to them, during the contact. Similarly, the elements of culture that will change depend on the group's acculturation strategies. At the individual level, both the behavior changes and acculturative stress phenomena are now known to be a function, at least to some extent, of

what people try to do during their acculturation, and the longer term outcomes (both psychological and sociocultural adaptations) often correspond to the strategic goals set by the groups of which they are members.

Four acculturation strategies have been derived from two basic issues facing all acculturating peoples. These issues are based on the distinction between orientations toward one's own group and those toward other groups. This distinction is rendered as a relative preference for maintaining one's heritage culture and identity and as a relative preference for having contact with and participating in the larger society along with other ethnocultural groups. This formulation is presented in Fig. 2.

These two issues can be responded to on attitudinal dimensions, represented by bipolar arrows. For purposes of presentation, generally positive or negative orientations to these issues intersect to define four acculturation strategies. These strategies carry different names, depending on which ethnocultural group (the dominant or nondominant one) is being considered. From the point of view of nondominant groups (on the left of Fig. 2), when individuals do not wish to maintain their cultural identity and instead seek daily interaction with other cultures, the assimilation strategy is defined. In contrast, when individuals place a value on holding onto their original culture while wishing to avoid interaction with others, the separation alternative is defined. When individuals have an interest in maintaining their original culture while in daily interactions with other groups, integration is the option. In this case, they maintain some degree of cultural integrity while seeking, as members of an ethnocultural group, to participate as an integral part of the larger social network. Finally, when individuals have little possibility of, or interest in, cultural maintenance (often for reasons of enforced cultural loss) and little interest in having relations with others (often for reasons of exclusion or discrimination), marginalization is defined.

This presentation was based on the assumption that nondominant groups and their individual members have the freedom to choose how they want to acculturate. This, of course, is not always the case. When the dominant group enforces certain forms of acculturation or constrains the choices of nondominant groups or individuals, other terms need to be used.

Integration can be "freely" chosen and successfully pursued by nondominant groups only when the dominant society is open and inclusive in its orientation

Issue 1:

Maintenance of culture and identity

Issue 2:

Relationships
sought
among
groups

Integration   Assimilation

Separation   Marginalization

Strategies of
nondominant groups

Multiculturalism   Melting pot

Segregation   Exclusion

Strategies of
dominant groups
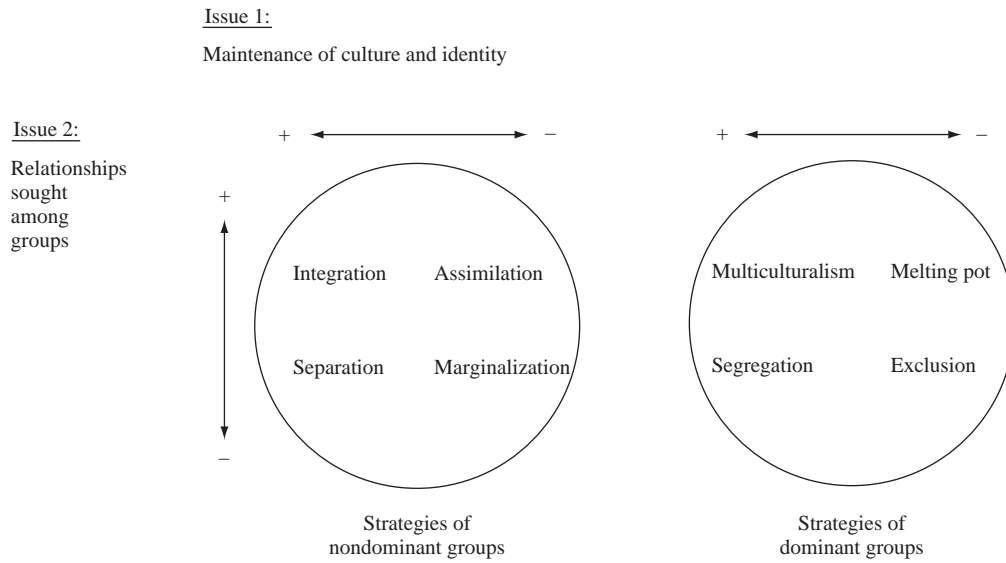
FIGURE 2   Four acculturation strategies based on two issues in ethnocultural groups and the larger society.

toward cultural diversity. Thus, a mutual accommodation is required for integration to be attained and involves the acceptance by both groups of the right of all groups to live as culturally different peoples. This strategy requires nondominant groups to adopt the basic values of the larger society, whereas the dominant group must be prepared to adapt national institutions (e.g., education, health, labor) to better meet the needs of all groups now living together in the plural society.

These two basic issues were initially approached from the point of view of the nondominant ethnocultural groups. However, the original anthropological definition clearly established that both groups in contact would become acculturated. Hence, a third dimension was added: the powerful role played by the dominant group in influencing the way in which mutual acculturation would take place. The addition of this third dimension produces the right side of Fig. 2. Assimilation, when sought by the dominant group, is termed the "melting pot"; however, when it is demanded by the dominant group, it is called the "pressure cooker." When separation is forced by the dominant group, it is "segregation." Marginalization, when imposed by the dominant group, is termed "exclusion." Finally, integration, when diversity is a feature of the society as a whole (including all of the various ethnocultural groups), is called "multiculturalism." With the use of this framework, comparisons can be made between individuals and their groups and between nondominant peoples and the larger society within which they

are acculturating. The ideologies and policies of the dominant group constitute an important element of ethnic relations research, whereas the preferences of nondominant peoples are a core feature in acculturation research. Inconsistencies and conflicts among these various acculturation preferences are sources of difficulty for acculturating individuals. In general, when acculturation experiences cause problems for acculturating individuals, we observe the phenomenon of acculturative stress.

4. ACCULTURATIVE STRESS

Three ways in which to conceptualize outcomes of acculturation have been proposed in the literature. In the first (behavioral shifts), we observe those changes in an individual's behavioral repertoire that take place rather easily and are usually nonproblematic. This process encompasses three subprocesses: culture shedding; culture learning; and culture conflict. The first two involve the selective, accidental, or deliberate loss of behaviors and their replacement by behaviors that allow the individual a better "fit" with the society of settlement. Most often, this process has been termed "adjustment" because virtually all of the adaptive changes take place in the acculturating individual, with few changes occurring among members of the larger society. These adjustments are typically made with minimal difficulty, in keeping with the appraisal

of the acculturation experiences as nonproblematic. However, some degree of conflict may occur, and this is usually resolved by the acculturating person yielding to the behavioral norms of the dominant group. In this case, assimilation is the most likely outcome.

When greater levels of conflict are experienced and the experiences are judged to be problematic but controllable and surmountable, the second approach (acculturative stress) is the appropriate conceptualization. In this case, individuals understand that they are facing problems resulting from intercultural contact that cannot be dealt with easily or quickly by simply adjusting or assimilating to them. Drawing on the broader stress and adaptation paradigms, this approach advocates the study of the process of how individuals deal with acculturative problems on first encountering them and over time. In this sense, acculturative stress is a stress reaction in response to life events that are rooted in the experience of acculturation.

A third approach (psychopathology) has had long use in clinical psychology and psychiatry. In this view, acculturation is nearly always seen as problematic; individuals usually require assistance to deal with virtually insurmountable stressors in their lives. However, contemporary evidence shows that most people deal with stressors and reestablish their lives rather well, with health, psychological, and social outcomes that approximate those of individuals in the larger society.

Instead of using the term "culture shock" to encompass these three approaches, the author prefers to use the term "acculturative stress" for two reasons. First, the notion of shock carries only negative connotations, whereas stress can vary from positive (eustress) to negative (dis-stress) in valence. Because acculturation has both positive (e.g., new opportunities) and negative (e.g., discrimination) aspects, the stress conceptualization better matches the range of affect experienced during acculturation. Moreover, shock has no cultural or psychological theory or research context associated with it, whereas stress (as noted previously) has a place in a well-developed theoretical matrix (i.e., stress–coping–adaptation). Second, the phenomena of interest have their life in the intersection of two cultures; they are intercultural, rather than cultural, in their origin. The term "culture" implies that only one culture is involved, whereas the term "acculturation" draws our attention to the fact that two cultures are interacting with, and producing, the phenomena. Hence, for both reasons, the author prefers the notion of acculturative stress to that of culture shock.

Relating these three approaches to acculturation strategies, some consistent empirical findings allow the following generalizations. For behavioral shifts, the fewest behavioral changes result from the separation strategy, whereas most result from the assimilation strategy. Integration involves the selective adoption of new behaviors from the larger society and retention of valued features of one's heritage culture. Marginalization is often associated with major heritage culture loss and the appearance of a number of dysfunctional and deviant behaviors (e.g., delinquency, substance abuse, familial abuse). For acculturative stress, there is a clear picture that the pursuit of integration is least stressful (at least where it is accommodated by the larger society), whereas marginalization is the most stressful. In between are the assimilation and separation strategies, with sometimes one and sometimes the other being the less stressful. This pattern of findings holds for various indicators of mental health.

5. ADAPTATION

As a result of attempts to cope with these acculturation changes, some long-term adaptations may be achieved. As mentioned earlier, adaptation refers to the relatively stable changes that take place in an individual or group in response to external demands. Moreover, adaptation may or may not improve the fit between individuals and their environments. Thus, it is not a term that necessarily implies that individuals or groups change to become more like their environments (i.e., adjustment by way of assimilation), but it may involve resistance and attempts to change their environments or to move away from them altogether (i.e., by separation). In this use, adaptation is an outcome that may or may not be positive in valence (i.e., meaning only well adapted). This bipolar sense of the concept of adaptation is used in the framework in Fig. 1 where long-term adaptation to acculturation is highly variable, ranging from well adapted to poorly adapted and varying from a situation where individuals can manage their new lives very well to one where they are unable to carry on in the new society.

Adaptation is also multifaceted. The initial distinction between psychological and sociocultural adaptation was proposed and validated by Ward in 1996. Psychological adaptation largely involves one's psychological and physical well-being, whereas sociocultural adaptation refers to how well an acculturating individual is able to manage daily life in the new cultural context. Although conceptually distinct, the two measures are empirically related to some extent (correlations between them are in the +.40 to +.50 range). However, they are

also empirically distinct in the sense that they usually have different time courses and different experiential predictors. Psychological problems often increase soon after contact, followed by a general (but variable) decrease over time. However, sociocultural adaptation typically has a linear improvement with time. Analyses of the factors affecting adaptation reveal a generally consistent pattern. Good psychological adaptation is predicted by personality variables, life change events, and social support, whereas good sociocultural adaptation is predicted by cultural knowledge, degree of contact, and positive intergroup attitudes.

Research relating adaptation to acculturation strategies allows for some further generalizations. For all three forms of adaptation, those who pursue and accomplish integration appear to be better adapted, whereas those who are marginalized are least well adapted. And again, the assimilation and separation strategies are associated with intermediate adaptation outcomes. Although there are occasional variations on this pattern, it is remarkably consistent and parallels the generalization regarding acculturative stress.

6. APPLICATIONS

There is now widespread evidence that most people who have experienced acculturation actually do survive. They are not destroyed or substantially diminished by it; rather, they find opportunities and achieve their goals, sometimes beyond their initial imaginings. The tendency to "pathologize" the acculturation process and outcomes may be due partly to the history of its study in psychiatry and in clinical psychology. Second, researchers often presume to know what acculturating individuals want and impose their own ideologies or personal views rather than informing themselves about culturally rooted individual preferences and differences. One key concept (but certainly not the only one) in understanding this variability, acculturation strategies, has been emphasized in this article.

The generalizations that have been made in this article on the basis of a wide range of empirical findings allow us to propose that public policies and programs that seek to reduce acculturative stress and to improve intercultural relationships should emphasize the integration approach to acculturation. This is equally true of national policies, institutional arrangements, and the goals of ethnocultural groups. It is also true of individuals in the larger society as well as members of nondominant acculturating groups.

In some countries, the integrationist perspective has become legislated in policies of multiculturalism that encourage and support the maintenance of valued features of all cultures while supporting full participation of all ethnocultural groups in the evolving institutions of the larger society. What seems certain is that cultural diversity and the resultant acculturation are here to stay in all countries. Finding a way in which to accommodate each other poses a challenge and an opportunity to social and cross-cultural psychologists everywhere. Diversity is a fact of contemporary life. Whether it is the "spice of life" or the main "irritant" is probably the central question that confronts us all—citizens and social scientists alike.

See Also the Following Articles

Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Cultural Complexity ■ Ideological Orientation and Values

Further Reading

Berry, J. W. (1976). *Human ecology and cognitive style: Comparative studies in cultural and psychological adaptation.* New York: Russell Sage/Halsted.

Berry, J. W. (1980). Acculturation as varieties of adaptation. In A. Padilla (Ed.), *Acculturation: Theory, models, and findings* (pp. 9–25). Boulder, CO: Westview.

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, *46*, 5–68.

Berry, J. W., Kalin, R., & Taylor, D. (1977). *Multiculturalism and ethnic attitudes in Canada.* Ottawa, Ontario: Supply & Services.

Berry, J. W., & Kim, U. (1988). Acculturation and mental health. In P. Dasen, J. W. Berry, & N. Sartorius (Eds.), *Health and cross-cultural psychology* (pp. 207–236). Newbury Park, CA: Sage.

Berry, J. W., Kim, U., Minde, T., & Mok, D. (1987). Comparative studies of acculturative stress. *International Migration Review*, *21*, 491–511.

Berry, J. W., Kim, U., Power, S., Young, M., & Bujaki, M. (1989). Acculturation attitudes in plural societies. *Applied Psychology*, *38*, 185–206.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications* (2nd ed.). New York: Cambridge University Press.

Graves, T. (1967). Psychological acculturation in a tri-ethnic community. *South-Western Journal of Anthropology*, *23*, 337–350.

Murphy, H. B. M. (1965). Migration and the major mental disorders. In M. B. Kantor (Ed.), *Mobility and mental health* (pp. 221–249). Springfield, IL: Thomas.

Ogbu, J. U. (1992). Understanding cultural diversity and learning. *Educational Researcher*, *21*, 5–14.

Redfield, R., Linton, R., & Herskovits, M. (1936). Memorandum on the study of acculturation. *American Anthropologist*, *38*, 149–152.

Richmond, A. (1993). Reactive migration: Sociological perspectives on refugee movements. *Journal of Refugee Studies*, *6*, 7–24.

Social Science Research Council. (1954). Acculturation: An exploratory formulation. *American Anthropologist*, *56*, 973–1002.

Ward, C. (1996). Acculturation. In D. Landis, & R. Bhagat (Eds.), *Handbook of intercultural training* (2nd ed., pp. 124–147). Newbury Park, CA: Sage.

Ward, C., Bochner, S., & Funham, A. (2001). *The psychology of culture shock*. London: Routledge.

Achievement Motivation in Academics

Dale H. Schunk

University of North Carolina, Greensboro, North Carolina, USA

1. Introduction
2. Background
3. Current Perspectives
4. Promoting Achievement Motivation
Further Reading

Achievement motivation is the desire to excel at effortful activities. Achievement motivation originally was thought of as extending across many areas, but current views conceive of it as more specific to situations. Personal, social, instructional, familial, and cultural factors affect achievement motivation, and parents and educators can help students to improve their achievement motivation.

GLOSSARY

achievement motivation The desire to excel at effortful activities.

attributions Perceived causes of outcomes.

classroom climate The social, psychological, and emotional characteristics of the classroom.

expectancy–value theory Theory predicting that achievement behavior depends on expectations about the outcomes of actions and how much those outcomes are valued.

goal What one is trying to accomplish.

outcome expectations Beliefs about the expected outcomes of one's actions.

peer networks Large groups of peers with whom students associate.

self-efficacy Beliefs about one's capabilities for learning or performing given actions.

Thematic Apperception Test (TAT) Projective personality measure in which people view ambiguous pictures and make up stories or answer questions.

value The importance of an outcome.

1. INTRODUCTION

At the end of her ninth-grade geometry class, Mrs. Lollar passed out the "Problem of the Week" for students to take home and solve. Two students in the class, Ashley and Marella, each have a B+ average. Because the problem of the week is worth extra credit, doing well on it could raise each of their averages to an A. That evening at home, Ashley looked at the problem for a while, spent a little time on it, but then quit without solving it. Marella studied the problem and began to work. When her parents called her to dinner, she was reluctant to come because she was working on it. After dinner, she worked on it some more and finally exclaimed, "Aha—I get it!" The next day in class, Marella was one of only three students to have solved the geometric proof. Ashley got some steps in the proof correct, but her solution was far from complete.

35

Achievement motivation is the desire to excel at effortful activities. In the opening scenario, it seems that Marella was more motivated to achieve in geometry than was Ashley. Marella displayed interest in solving the proof, persisted at it, was excited when she understood what to do, and solved it correctly.

Achievement motivation has a long history in psychology and education, and for good reason. Motivation to achieve is necessary for all but the simplest tasks. Achievement motivation helps students to learn in school, fuels creative activities, and helps individuals and societies to attain goals.

Despite the intuitive importance of achievement motivation, researchers disagree on its critical components. The next two sections examine the background of achievement motivation and some current perspectives.

2. BACKGROUND

It is difficult to pinpoint the historical onset of the study of achievement motivation because the human desire to achieve has been of interest for ages. The scientific study of achievement motivation received impetus from work by Murray, who included it as one of several human needs that contribute to personality development. Murray also devised the Thematic Apperception Test (TAT) to study personality processes. The TAT is a projective technique in which people view a series of ambiguous pictures (i.e., inkblots) and make up a story or answer questions for each picture.

McClelland and colleagues adapted the TAT to study achievement motivation. People were shown pictures of individuals in ambiguous situations (e.g., a student at a desk holding a pencil and looking into the air) and were asked questions such as the following: "What is happening?," "What led up to this situation?," "What is wanted?," and "What will happen?" Responses were scored and categorized according to strength of achievement motivation. Unfortunately the TAT has some measurement problems, and TAT achievement motivation scores often did not relate well to other measures of achievement. Over the years, researchers have devised other methods for assessing achievement motivation.

Important early work on achievement motivation was done by Atkinson and colleagues. Atkinson drew on work by Lewin and others on the level of aspiration or the goal that people set in a task. Lewin's research showed that successes raised and failures lowered the level of aspiration, that people felt more successful when they met the goals they had set for themselves than when they attained objective standards, and that the level of aspiration reflected individual and group differences.

Atkinson's expectancy–value theory of achievement motivation states that behavior depends on how much people value a particular goal and their expectancy of attaining that goal as a result of performing in a given fashion. Atkinson postulated that achievement behavior involved a conflict between a motive to approach a task (hope for success) and a motive to avoid it (fear of failure). These motives conflict because any achievement task carries with it the possibilities of success and failure. The achievement motivation that results in any situation depends on people's expectancies of success and failure and their incentive values of success and failure.

The historical research focused on achievement motivation globally, that is, motivation across many situations. But research and everyday observations show that people rarely are motivated to achieve at high levels in every situation. Rather, students typically have greater achievement motivation in some content areas than in others. In the opening scenario, although Marella's achievement motivation was higher than Ashley's in geometry, Ashley may strive to excel more than Marella in history. Given that the achievement motive differs depending on the domain (and even depending on the task within the domain), the validity of general achievement motivation is questionable.

3. CURRENT PERSPECTIVES

People who study achievement motivation today believe that it is more situationally specific and is affected by many factors: personal, social, instructional, familial, and cultural (Table I). To understand why students differ in achievement motivation, one must examine the roles of these factors in their lives. Although researchers agree that achievement motivation is complex, they disagree on which factors are the most important.

Personal factors reside within students. Some key personal factors are goals, expectations, values, and attributions. Goals are what one is trying to accomplish. Goals usually are cast in terms of products—Marella and Ashley's goal was to work the geometry proof—but they also can represent processes. Thus, an academic goal may be to improve one's skill in comprehending scientific texts. Goals contribute to achievement motivation because people pursuing goals persist and expend effort to succeed. As they work on the tasks, they evaluate their goal processes, and the belief that they are making progress sustains motivation.

TABLE I
Influences on Achievement Motivation

Factors	Sample influences
Personal	Goals, outcome expectations, self-efficacy, values, attributions
Social	Peers, peer models, peer networks
Instructional	Planning, decision making, instructional grouping, interactions, activities, classroom climate
Familial	Socioeconomic status, capital, guidance
Cultural	Educational emphasis, social orientation

But goals actually are less important than their properties: specificity, proximity, and difficulty. Goals may be specific (e.g., read 10 pages) or general (e.g., read some pages), proximal (e.g., read 10 pages tonight) or distant (e.g., read 10 pages by next week), and difficult (e.g., read 400 pages) or easier (e.g., read 50 pages). From a motivational perspective, goals that are specific, proximal, and moderately difficult produce higher achievement motivation than do goals that are general, distant, and either too difficult or too easy. Thus, motivation is not aided when goals denote general outcomes (because nearly any action will satisfy them), are temporally distant (it is easy to put off until tomorrow what does not need to be done today), and too difficult or too easy (people are not motivated to attempt the impossible and may procrastinate completing easy tasks).

Expectations can be of two types. Outcome expectations refer to the expected outcomes of one's actions (e.g., "If I study hard, I should make a high grade on the test," "No matter how hard I study, I probably will fail the test"). Outcome expectations motivate students because the belief that a certain action will lead to a given outcome should lead students to pursue or avoid that action, assuming that they value that outcome or want to avoid it.

A second type of expectation involves beliefs about one's capabilities to learn or perform given actions. Bandura and other researchers termed this "self-efficacy." Self-efficacious students believe that they can study diligently, whereas those lacking efficacy might believe that they cannot. Marella's self-efficacy in geometry likely was higher than Ashley's. Self-efficacy is important because students who believe that they can learn or perform desired actions are more likely to choose to engage in them, expend effort, and persist.

Self-efficacy and outcome expectations often are related, but they need not be. Students who believe that they are capable typically expect to perform well and receive high grades and other rewards. However,

students may believe that diligent studying will lead to good grades (i.e., positive outcome expectation) but also may doubt their ability to study diligently (i.e., low self-efficacy), in which case they might feel demoralized. Positive outcome expectations and strong self-efficacy for learning and performing capably produce high achievement motivation.

The values and/or importance that students ascribe to learning and achievement are central components of motivation. Those who do not value what they are learning are not motivated to improve or perform well. Research shows that value relates positively to persistence, choice, and performance. Students who value learning choose challenging activities, persist at them, and perform well. Thus, students who value history are apt to study diligently for tests, set goals for their learning, monitor their learning progress, use effective learning strategies, and not be daunted by obstacles. Marella likely valued geometry more than did Ashley.

Attributions are perceived causes of outcomes or the factors that people believe are responsible for their successes and failures. In achievement situations, learners may ask themselves questions such as the following: "Why did I get an A on my biology test?" and "Why can't I learn French?"

Weiner and colleagues formulated an attribution theory of achievement motivation that contends that each attribution can be classified along three dimensions: locus, stability, and controllability. The locus dimension refers to whether the attribution is internal or external to the person. The stability dimension denotes how much the attribution varies over time. The controllability dimension involves the extent that the attribution is under the individual's control.

Weiner found that common attributions in achievement settings are ability, luck, task difficulty, and effort; however, there are many others (e.g., fatigue, illness, personal dislike, time available, attitude). Ability is internal, stable, and uncontrollable (although one's ability

can improve over time); luck is external, unstable, and uncontrollable; task difficulty is external, stable (assuming that the task does not change), and uncontrollable; and effort is internal, unstable (although a general effort factor also seems to exist), and controllable.

Students who attribute academic success to high ability and effort are more likely to expect future success than are those who attribute it to task ease and good luck. Students who attribute failure to low ability are less likely to expect future success than are those who attribute it to low effort. The latter finding is especially critical because a low expectation of future success stifles motivation. From the perspective of achievement motivation, it is far better to stress effort than ability because the former has stronger motivational effects. Of course, as skills develop, students do become more able, so the attribution of success to ability is credible. Teachers often have difficulty in motivating students who believe that they lack the ability to succeed and that no amount of effort will help them to be successful.

Social factors are those that are inherent in interactions with others. Peers are a key social group for students. Peer models are especially influential, especially those who are similar to observers in important ways, because they convey that tasks are valuable.

Peer group goals are highly valued by students. Thus, students may want to be liked and approved, develop social and/or intimate relationships, learn to cooperate, win favors, and be sensitive to the needs of others. Students' perceptions of their capabilities also are affected by peers through social comparisons. Observing similar peers succeeding can raise observers' self-efficacy.

Peer networks are large groups of peers with whom students associate. Students in networks tend to be highly similar and, thus, are key models. Students' motivational engagement in school across the school year is predicted by their network membership at the start of the year. Those in more academically inclined networks demonstrate higher academic motivation than do those in low-motivation networks.

Although students may select their peer networks, parents can play a key role by "launching" their children onto particular trajectories. For example, parents who want their children to be academically oriented are likely to involve them in activities and groups that stress academics. Peers with whom children associate reinforce the emphasis on academics.

Instructional factors include teacher planning and decision making, grouping for instruction, teacher–student interactions (including feedback to students), activities, and classroom climate. Teachers can enhance student motivation by planning interesting activities that maximize student involvement in lessons. Teachers who plan only lectures are less apt to promote student motivation. Teachers also can promote motivation by basing instructional decisions not only on how well students are learning but also on how much the material appeals to students.

Three types of grouping structures are competitive, cooperative, and individualist. Competitive situations are those in which the goals of individuals are linked negatively such that if one attains his or her goal, the chance that others will attain their goals diminishes. Cooperative structures are those in which the goals of the group members are linked positively such that one can attain his or her goal only if others attain their goals. In individualist situations, there is no link among the goals of individuals such that one's goal attainment has no effect on the goal attainment of others. From a motivational perspective, competitive situations highlight differences among students, and lower achievers may become discouraged if they believe that they have no chance to earn rewards. Cooperative situations are better if all students contribute to the project. If only one or two students do most of the work, there is apt to be resentment. In individualist situations, achievement motivation can be developed and sustained when learners focus on their progress or on how much better they are performing now compared with earlier.

Performance feedback provides information on accuracy of work and may include corrective information (e.g., "The first part is correct, but you need to bring down the next number"). Motivational feedback can provide information on progress and competence (e.g., "You've gotten much better at this"), link student performance with one or more attributions (e.g., "You've been working hard"), and inform students about how well they are applying a strategy and how strategy use is improving their work (e.g., "You got it correct because you followed the proper method"). Feedback motivates students when it informs them that they are making progress and becoming more competent.

Classroom climate refers to the atmosphere of the classroom—its social, psychological, and emotional characteristics. Climate is important for motivation because classroom interactions define the climate. Climate often is referred to in terms such as "warm," "cold," "permissive," "democratic," and "autocratic." Research shows that a democratic environment—one based on mutual respect and collaboration—fosters goal attainment by learners without their becoming frustrated or aggressive.

There are many familial factors, but a key one is socioeconomic status, whose link with students'

academic motivation is well established. Children from lower socioeconomic backgrounds typically display lower motivation and achievement, and are at greater risk for school failure and dropout, compared with children from wealthier families. But socioeconomic status is only a descriptive term and does not explain why effects occur. There also are many people who grew up in impoverished environments but became well educated and successful.

A major contributor is a family's capital or resources. Poor families have less financial resources to support their children's learning outside of school than do wealthier families. Socialization in lower class homes often does not match or prepare students for the middle-class orientation of schools and classrooms. Lower socioeconomic students might not understand the full benefits of schooling or comprehend that getting a good education will increase their chances of college acceptance, good jobs, and financial stability. Because of family financial strain, they also might not be able to resist the short-term benefits of working in favor of the long-term benefits of schooling. Such students might have few, if any, educated role models in their environments. Many enter school without the needed social, cognitive, and emotional prerequisites to learn successfully.

Families continue to influence children's motivation throughout childhood and adolescence by steering children in given directions. Families that provide rich resources in the home and guide their children into activities that stress motivation and achievement are apt to develop children higher in achievement motivation.

Cultural factors also affect motivation. Cross-cultural research shows that there are differences in how much cultures emphasize education and motivation for learning. For example, children from Asian cultures often place greater emphasis on effort as a cause of success than do students in the United States. Some cultures are more socially oriented, whereas others place more emphasis on individual accomplishments. Although it is hard to make generalizations about cultures given that not all members of a culture act alike, it does seem that motivational differences stem in part from influences in cultural background.

4. PROMOTING ACHIEVEMENT MOTIVATION

How can parents and educators help to improve students' achievement motivation for academics? First, they can help students to set challenging but attainable goals. Students must believe that the goals are attainable; they will not be motivated to attempt what they see as impossible. Teachers and parents might need to work with students to ensure that goals are realistic; if students are unaware of the demands of an assignment, they might set an unattainable goal. Challenging and attainable goals help to build self-efficacy as students perceive that they are making progress.

Second, it is important for parents and educators to stress the value of learning. Motivation is enhanced when learners understand how they can use what they are learning and how it will help them in the future. People are not motivated to engage in "busy work." Marella and Ashley's teacher would do well to point out applications of geometry in daily life.

Third, parents and educators should build students' perceptions of their competence or self-efficacy. Self-efficacy is increased when learners believe that they are developing skills, making progress toward their goals, and performing better. Self-efficacy develops through actual performance accomplishments, vicarious (modeled) experiences, and social persuasion. Teachers and parents should ensure that students work at tasks on which they can be successful, observe similar peers succeeding, and receive feedback indicating that they are capable (e.g., "I know you can do this").

Fourth, parents should get involved in children's learning. Parents can be involved in school by assisting in classrooms and at home. Parents serve as models for learning, and they can help to establish a productive study environment. They also can be taught tutoring skills. The importance of parents as motivational models cannot be overemphasized.

Fifth, teachers should use peers effectively to build motivation. Teachers who group students for collaborative project work must ensure that each student has responsibility for part of the task so that the bulk of the work is not done by one or two students. In using models for comparisons, it is imperative that students view the models as similar to themselves. Motivation of lower achieving students might not improve when these students are asked to observe higher achieving students perform.

Finally, parents and educators should use feedback to teach and motivate. Feedback typically informs students whether they are correct or incorrect and what to do to perform better. Feedback motivates when it informs students about their progress—how much better they are performing now than they were previously—and when it links their performances to effort, good strategy use, and enhanced ability.

Academic motivation for academics is important not only for schooling but also for the future. Educators want students to continue to be motivated to learn once they leave their classrooms. This is the essence of "lifelong learning" by which citizens move their societies forward.

See Also the Following Articles

Educational Achievement and Culture ■ Learning ■ Learning Styles and Approaches to Studying ■ Motivation and Culture

Further Reading

Ames, C. (1984). Competitive, cooperative, and individualistic goal structures: A cognitive–motivational analysis. In R. Ames, & C. Ames (Eds.), *Research on motivation in education* (Vol. 1, pp. 177–207). New York: Academic Press.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 359–372.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371–399.

McClelland, D., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive.* New York: Appleton–Century–Crofts.

Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Steinberg, L., Brown, B. B., & Dornbusch, S. M. (1996). *Beyond the classroom: Why school reform has failed and what parents need to do.* New York: Simon & Schuster.

Weiner, B. (1992). *Human motivation: Metaphors, theories, and research*. Newbury Park, CA: Sage.

Achievement Tests

Gregory J. Cizek

University of North Carolina, Chapel Hill, North Carolina, USA

GLOSSARY

ability test A procedure designed to measure an examinee's potential for achievement, sometimes also referred to as an aptitude test.

achievement test A procedure designed to measure an examinee's attainment of knowledge or skills.

battery A collection of related subtests, the performance on which can be aggregated to form a composite or "total battery" score on the entire test.

classroom assessments Achievement tests that are developed, administered, and scored for classroom use.

content standards Statements that describe specific knowledge or skills over which examinees are expected to have mastery for a given age/grade and subject area.

criterion-referenced test (CRT) An instrument whose primary purpose is to gauge whether an examinee knows or can do specific things.

high-stakes/low-stakes Terms referring to the severity of the consequences associated with test performance.

inference A reasoned leap, supported by evidence, made whenever observed test scores are interpreted in terms of examinees' underlying standing on a construct or actual level of knowledge or skill.

norm-referenced test (NRT) An instrument whose primary purpose is to describe the relative standing of examinees at a particular age or grade level.

performance standards Prescribed levels of performance representing differential degrees of knowledge or skill with respect to a set of content standards.

reliability The dependability of test scores; because an assignment or test consists only of a sample of questions or tasks, and because both the students who respond and those who score the responses are susceptible to various unpredictabilities in their performance (called random errors), no score can considered to be a perfectly dependable representation of a student's performance.

standardized A term used to describe any instrument that is developed, administered, and scored under controlled conditions.

standards-referenced test (SRT) An instrument whose primary purpose is to describe examinees' performance with respect to a set of content standards.

validity The degree to which the conclusions yielded by any sample of behavior (e.g., test, quiz, interview) are meaningful, accurate, and useful; in the context of achievement testing, validity is the degree to which students' test scores yield inferences about the students' level of knowledge, skill, or ability that are "correct" or "on target."

Achievement tests are used in diverse contexts to measure the degree to which examinees can demonstrate acquisition of knowledge or skills deemed to be important. The contexts range from teacher-made testing in elementary and secondary school settings to

high-stakes testing for college admission, licensure to practice a profession, or certification. The design of achievement tests varies depending on whether the inference intended to be drawn regarding examinees' performance is the absolute or relative level of mastery of specific knowledge and skills.

1. INTRODUCTION

Achievement testing is a general term used to describe any measurement process or instrument whose purpose is to estimate an examinee's degree of attainment of specified knowledge or skills. Beyond that central purpose, achievement tests differ according to their specific intended inference. Common inferences include either absolute level of performance on the specified content or relative standing vis-à-vis other examinees on the same content. Achievement tests may be group or individually administered. They may consist of differing formats, including multiple-choice items, essays, performance tasks, and portfolios.

Achievement tests are administered in diverse contexts. For example, they are used when the school-related skills of preschool pupils are measured to assess their readiness for kindergarten. During the K–12 school years, students typically take a variety of achievement tests, ranging from teacher-made informal assessments, to commercially prepared achievement batteries, to state-mandated high school graduation tests. Following formal schooling, achievement tests are administered to assess whether examinees have an acceptable level of knowledge or skill for safe and competent practice in a regulated profession for which licensure is required. In other situations, professional organizations establish certification procedures, often including an achievement test, to determine examinees' eligibility to attain a credential or certification of advanced status in a field. Even the ubiquitous requirements to obtain a driver's license involve an achievement testing component.

Although the purposes and contexts may vary, fairly uniform procedures are implemented for developing achievement tests and for evaluating their technical quality. Several sources exist for potential users of achievement tests to ascertain the quality of a particular test and its suitability for their purposes.

In any context where an achievement test is used, consequences for individual persons or groups may follow from test performance. In addition, the context and extent of achievement testing may have broad and sometimes unforeseen consequences affecting, for example, the security of tests, the formats used for testing, and the relationship between testing and instruction.

2. DEFINITIONS AND EXAMPLES

Achievement testing refers to any procedure or instrument that is used to measure an examinee's attainment of knowledge or skills. Achievement testing can be done informally, as in when a teacher asks a student to perform a skill such as reading aloud or demonstrating correct laboratory technique. More formal, and perhaps more common, achievement tests are routinely administered in educational and occupational settings. Examples of formal achievement testing in education would include spelling tests, chemistry lab reports, end-of-unit tests, homework assignments, and so on.

More formal achievement testing is evident in large-scale, commercially available standardized instruments. The majority of these achievement tests would be referred to as standardized to the extent that the publishers of the instruments develop, administer, and score the tests under uniform, controlled conditions. It is important to note, however, that the term "standardized" is (a) unrelated to test format (although the multiple-choice format is often used for standardized tests, any format may be included) and (b) not synonymous with norm referenced (although sometimes the term "standardized" is used to indicate that a test has norms).

Examples of standardized achievement tests used in K–12 education would include the Iowa Tests of Basic Skills, the 10th edition of the Stanford Achievement Test, and the TerraNova. These tests ordinarily consist of several subtests, measuring achievement in specific narrow areas such as language arts, mathematics, science, and study skills. The composite index formed from these subtests (often referred to as a "complete battery score") provides a more global measure of academic achievement.

The preceding tests are also usually administered in a group setting, although individually administered achievement tests are also available and are designed for administration in a one-on-one setting with individual students, usually of very young age. Examples of individually administered achievement tests include the Woodcock–Johnson III Tests of Achievement, the third edition of the Developmental Indicators for the Assessment of Learning, and the Brigance Comprehensive Inventory of Basic Skills.

Following secondary schooling, achievement testing continues in colleges and universities, primarily in the

form of classroom achievement measures, but would also include standardized in-training examinations and board examinations for persons pursuing professional careers. Achievement testing has a long history in diverse occupational fields. Achievement tests are routinely administered to ascertain levels of knowledge or skills when screening or selecting applicants for positions in business and industry. These tests have traditionally been administered in paper-and-pencil format, although technology has enabled administration via computer or over the Internet to be secure, fast, and accessible. For example, one vendor of computerized achievement tests offers computerized "work sample" achievement tests to assist human resources personnel in selecting applicants for positions in legal offices, food service, information technology, accounting, medical offices, and others. Many state, federal, and private organizations also provide achievement tests for a variety fields in which licensure or certification is required.

3. TYPES OF ACHIEVEMENT TESTS

In the previous section, it was noted that achievement tests could be categorized according to administration (group or individual) and scale (informal classroom tests or more formal commercially available tests). Another more important distinction focuses on the intended purpose, use, or inference that is to be made from the observed test score.

Less formal classroom achievement tests are usually developed by a teacher to align with an instructional unit, or they may be pre-prepared by publishers of classroom textbooks or related materials. The primary purposes of such tests are for educators' use in refining instruction and assigning grades as well as for both teacher and pupil use in understanding and responding to individual students' strengths and weaknesses.

More formal standardized achievement tests can also be categorized according to the inferences they yield. Three such types of tests—criterion-referenced tests (CRTs), standards-referenced tests (SRTs), and norm-referenced tests (NRTs)—are described in this section.

CRTs are designed to measure absolute achievement of fixed objectives comprising a domain of interest. The content of CRTs is narrow, highly specific, and tightly linked to the specific objectives. Importantly, a criterion for judging success on a CRT is specified a priori, and performance is usually reported in terms of pass/fail, number of objectives mastered, or similar terms. Thus, an examinee's performance or score on a CRT is interpreted with reference to the criterion. The written driver's license test is a familiar example of a CRT.

SRTs are similar to CRTs in that they are designed to measure an examinee's absolute level of achievement vis-à-vis fixed outcomes. These outcomes are narrowly defined and are referred to as content standards. Unlike CRTs, however, interpretation of examinees' performance is referenced not to a single criterion but rather to descriptions of multiple levels of achievement called performance standards that illustrate what performance at the various levels means. Typical reporting methods for SRTs would consist of performance standard categories such as basic, proficient, and advanced or beginner, novice, intermediate, and expert levels. Familiar examples of SRTs include state-mandated testing for K–12 students in English language arts, mathematics, and so on to the extent that the tests are aligned with the state's content standards in those subjects. At the national level, the National Assessment of Educational Progress (NAEP) is administered at regular intervals to samples of students across the United States.

NRTs are designed to measure achievement in a relative sense. Although NRTs are also constructed based on a fixed set of objectives, the domain covered by an NRT is usually broader than that covered by a CRT. The important distinction of NRTs is that examinee performance is reported with respect to the performance of one or more comparison groups of other examinees. These comparison groups are called norm groups. Tables showing the correspondence between a student's performance and the norm group's performance are called norms. Thus, an examinee's performance or score on an NRT is interpreted with reference to the norms. Typical reporting methods for NRTs include z scores, percentile ranks, normal curve equivalent scores, grade- or age-equivalent scores, stanines, and other derived scale scores. Familiar examples of NRTs include the Iowa Tests of Basic Skills (ITBS), the Scholastic Assessment Test (SAT), and the Graduate Record Examinations (GRE).

Many publishers of large-scale achievement tests for school students also provide companion ability tests to be administered in conjunction with the achievement batteries. The tests are administered in tandem to derive ability/achievement comparisons that describe the extent to which a student is "underachieving" or "overachieving" in school given his or her measured potential. Examples of these test pairings include the Otis–Lennon School Abilities Test (administered with the Stanford Achievement Test) and the Cognitive Abilities Test (administered with the ITBS).

4. ACHIEVEMENT TEST CONSTRUCTION

Rigorous achievement test development consists of numerous common steps. Achievement test construction differs slightly based on whether the focus of the assessment is classroom use or larger scale. Table I provides a sequence listing 18 steps that would be common to most achievement test development.

In both large and smaller contexts, the test maker would begin with specification of a clear purpose for the test or battery and a careful delineation of the domain to be sampled. Following this, the specific standards or objectives to be tested are developed. If it is a classroom achievement test, the objectives may be derived from a textbook, an instructional unit, a school district curriculum guide, content standards, or another source. Larger scale achievement tests (e.g., state mandated, standards referenced) would begin the test development process with reference to adopted state content standards. Standardized norm-referenced instruments would ordinarily be based on large-scale curriculum reviews, based on analysis of content standards adopted in various states, or

TABLE I
Common Steps in Achievement Test Development

1. Establish need, purpose
2. Delineate domain to be tested
3. Develop specific objectives, content standards
4. Decide on item and test specifications, formats, length, costs
5. Develop items, tasks, scoring guides
6. Conduct item/task review (editorial, appropriateness, alignment, sensitivity)
7. Pilot/Field test items/tasks/scoring guides
8. Review item/task performance
9. Create item bank/pool
10. Assemble test form(s) according to specifications
11. Develop test administration guidelines, materials
12. Establish performance standards
13. Administer operational test forms
14. Score test
15. Evaluate preliminary item/task and examinee performance
16. Report scores to appropriate audiences
17. Evaluate test, document test cycle
18. Update item pool, revise development procedures, develop new items/tasks

promulgated by content area professional associations. Licensure, certification, or other credentialing tests would seek a foundation in job analysis or survey of practitioners in the particular occupation. Regardless of the context, these first steps involving grounding of the test in content standards, curriculum, or professional practice provide an important foundation for the validity of eventual test score interpretations.

Common next steps would include deciding on and developing appropriate items or tasks and related scoring guides to be field tested prior to actual administration of the test. At this stage, test developers pay particular attention to characteristics of items and tasks (e.g., clarity, discriminating power, amenability to dependable scoring) that will promote reliability of eventual scores obtained by examinees on the operational test.

Following item/task tryout in field testing, a database of acceptable items or tasks, called an item bank or item pool, would be created. From this pool, operational test forms would be drawn to match previously decided test specifications. Additional steps would be required, depending on whether the test is to be administered via paper-and-pencil format or computer. Ancillary materials, such as administrator guides and examinee information materials, would also be produced and distributed in advance of test administration. Following test administration, an evaluation of testing procedures and test item/task performance would be conducted. If obtaining scores on the current test form that were comparable to scores from a previous test administration is required, then statistical procedures for equating the two test forms would take place. Once quality assurance procedures have ensured accuracy of test results, scores for examinees would be reported to individual test takers and other groups as appropriate. Finally, documentation of the entire process would be gathered and refinements would be made prior to cycling back through the steps to develop subsequent test forms (Steps 5–18).

5. EVALUATING ACHIEVEMENT TESTS

In some contexts, a specific achievement test may be required for use (e.g., state-mandated SRTs). However, in many other contexts, potential users of an achievement test may have a large number of options from which to choose. In such cases, users should be aware of the aids that exist to assist them in making informed choices.

One source of information about achievement tests is the various test publishers. Many publishers have online information available to help users gain a better understanding of the purposes, audiences, and uses of their products. Often, online information is somewhat limited and rather nontechnical. However, in addition to providing online information, many publishers will provide samples of test materials and technical documentation on request to potential users. Frequently, publishers will provide one set of these packets of information, called specimen sets, at no charge for evaluation purposes.

When evaluating an achievement test, it is important to examine many aspects. A number of authorities have provided advice on how to conduct such a review. For example, one textbook for school counselors by Whiston contains a section titled "Selection of an Assessment Instrument" that consists of several pages of advice and a user-friendly checklist. The single authoritative source for such information would likely be the Standards for Educational and Psychological Testing, jointly sponsored by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education.

Finally, a particularly useful framework for evaluating achievement tests was developed by Rudner in 1994. Table II provides a modified version of key points identified by Rudner that should be addressed when choosing an achievement test.

It is likely that some potential users will not have the time or technical expertise necessary to fully evaluate an achievement test independently. A rich source of information exists for such users in the form of published reviews of tests. Two compilations of test reviews are noteworthy: Mental Measurements Yearbook (MMY) and Tests in Print. These references are available in nearly all academic libraries. In the case of MMY, the editors of these volumes routinely gather test materials and forward those materials to two independent reviewers. The reviewers provide brief (two- to four-page) summaries of the purpose, technical qualities, and administration notes for the test. Along with the summaries, each entry in MMY contains the publication date for the test, information on how to contact the publisher, and cost information for purchasing the test. In these volumes, users can compare several options for an intended use in a relatively short time.

TABLE II
Evaluation Criteria for Achievement Tests

1. Is the purpose of the test clearly stated? What achievement construct is claimed? Is the construct or intended content domain clearly delineated?
2. What are the intended uses of the test? What are the intended audiences for the test results?
3. For what ages, grade levels, or subject areas is the test intended?
4. Are the test materials (e.g., booklets, answer sheets) clear, engaging, and appropriate for the age/grade level of the examinees?
5. What are the costs of the test materials, scoring, training personnel, and time required to administer the test?
6. Are the procedures for administering the test clear? Is the information provided sufficiently detailed so as to provide consistent administrations across users and contexts?
7. What are the qualifications of those who participated in the development of the test? What qualifications are required for test administrators?
8. How were samples selected for developing, pilot testing, norming, estimating reliability, screening out potentially biased items, and gathering validity evidence for the test? Were the samples relevant? Were they representative? Were they collected recently?
9. Does the test yield scores of acceptable reliability? Were appropriate methods used to compute reliability estimates? If decisions are to be made based in part on test performance, what is the evidence regarding decision consistency?
10. Is there adequate validity evidence to support the intended inferences, uses, interpretations, or meanings that will be made from test scores? Is there evidence to support the use of the test with various groups (including non-native speakers of English, special needs learners, students in need of testing accommodations, etc.)?
11. Is the scoring system likely to produce accurate scores? If hand scoring is involved, are scoring keys easy to use and is conversion of raw scores to derived scores facilitated by tables, look-up charts, and the like? If machine scoring is used, are answer documents easy for examinees to use? Are timelines and costs for scoring reasonable?
12. Do score reports provide a clear, detailed, and comprehensible summary of performance and diagnostic information? Are users appropriately cautioned about likely misinterpretations?

Source. Adapted from Rudner (1994).

A fee-based search capability for locating test reviews is available at the MMY Web site (www.unl.edu).

See Also the Following Articles

Educational Achievement and Culture

Further Reading

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Brigance, A. H., & Glascoe, F. P. (1999). *Brigance Comprehensive Inventory of Basic Skills* (rev. ed.). North Billerica, MA: Curriculum Associates.

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G. D. Phye (Ed.), *Handbook of classroom assessment* (pp. 1–32). San Diego: Academic Press.

Cizek, G. J. (2003). *Detecting and preventing classroom cheating: Promoting integrity in schools.* Thousand Oaks, CA: Corwin.

CTB/McGraw–Hill. (1997). *TerraNova.* Monterey, CA: Author.

Gronlund, N. E. (1993). *How to make achievement tests and assessments.* Boston: Allyn & Bacon.

Harcourt Educational Measurement. (2002). *Stanford Achievement Test* (10th ed.). San Antonio, TX: Author.

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *Iowa Tests of Basic Skills.* Itasca, IL: Riverside.

Mardell-Czudnowski, C., & Goldenberg, D. S. (1998). *Developmental indicators for the assessment of learning* (3rd ed.). Circle Pines, MN: American Guidance Services.

Rudner, L. (1994, April). *Questions to ask when evaluating tests* (ERIC/AE Digest, EDO-TM-94-06). Washington, DC: ERIC Clearinghouse on Assessment and Evaluation.

Whiston, S. C. (2000). *Principles and applications of assessment in counseling.* Belmont, CA: Wadsworth.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III Tests of Achievement.* Itasca, IL: Riverside.

Advertising and Culture

Sharon Shavitt and Jing Zhang

University of Illinois at Urbana–Champaign, Champaign, Illinois, USA

GLOSSARY

advertising appeal The central theme or idea behind an advertising message.

collectivism A cultural value system that emphasizes in-group goals such as family integrity, harmonious relationships, and the well-being of the in-group.

horizontal–vertical dimension A cultural distinction within the individualism–collectivism framework; members of horizontal societies value equality and view the self as having the same status as others in society, whereas members of vertical societies view the self as differing from others along a hierarchy, accepting inequality, and believing that rank has its privileges.

independent self-construal A view of the self that focuses on one's own psychological traits, feelings, and actions.

individualism A cultural value system that emphasizes being independent and pursuing individual goals rather than in-group goals.

interdependent self-construal A view of the self that focuses on one's social roles or memberships.

Culture is a crucial concept to the understanding of advertising and consumer behavior because it is the lens through which people view advertising messages and products. This article reviews and discusses a series of studies on the relations between culture and advertising. The content is organized around different levels of analysis of culture, for example, national level versus individual level. The individualism–collectivism cultural construct and independent–interdependent self-construal construct are given special attention because extensive research has demonstrated the persuasion implications of these variables for advertising content, attitude toward advertisements, and cognitive processing of advertisements.

1. ADVERTISING AND ADVERTISING APPEALS

Advertising has been defined as the nonpersonal communication of information, usually paid for and usually persuasive in nature, about products (goods or services) or ideas by identified sponsors through various media. The central theme or idea behind an advertising message is called the advertising appeal.

2. APPROACHES TO CULTURE IN THE CONTEXT OF ADVERTISING AND CONSUMER BEHAVIOR

Intensified global marketing and advertising activities have given rise in the advertising industry to the

advertising standardization versus specification debate. The standardization proponents believe that people around the world have the same needs and desires and that they would react to advertising messages in the same or a similar way. The specification proponents argue that cultural differences among nations make advertising one of the most difficult marketing elements to standardize.

Culture consists of shared elements that provide the standards for perceiving, believing, evaluating, communicating, and acting among those who share a language, a historical period, and a geographical location. As a psychological construct, culture can be studied at multiple levels—across nations, across individuals within nations, and even within individuals through the priming of cultural values. Regardless of how culture is studied, cultural distinctions have been demonstrated to have important implications for advertising content, advertising persuasion, and consumer response.

2.1. Cultural Differences at the National Level

2.1.1. Individualism and Collectivism

Individualism and collectivism comprise the main cultural distinction that has been explored in studying both advertising content and the persuasiveness of ad appeals. Individualism and collectivism have been conceptualized as two powerful cultural models that represent broad differences among nations. As Hofstede, Triandis, and others have proposed, members of collectivistic cultures endorse in-group goals, such as family integrity, harmonious relationships, and the well-being of the in-group, whereas members of individualistic cultures endorse being independent and pursuing individual goals instead of in-group goals. Extensive cross-cultural (cross-national) data have shown that North American and most European countries, such as the United States, Canada, Germany, and Denmark, are individualistic societies and that most East Asian and Latin American countries, such as China, Korea, Japan, and Mexico, are collectivistic societies.

Past research has indicated that more individualistic and less collectivistic advertising appeals are present in individualistic cultures than in collectivistic cultures. For example, Han and Shavitt's 1994 study found that Korean ads tended to employ appeals emphasizing in-group benefits, harmony, and family integrity, whereas U.S. ads tended to employ appeals emphasizing individual enjoyment, personal success, and independence.

Similarly, in other research, more conformity themes (e.g., respect for collective values and beliefs) and fewer uniqueness themes (e.g., rebelling against collective values and beliefs) were found to be present in Korean ads compared with U.S. ads. In studying humorous appeals, ads from both Korea and Thailand were found to contain more group-oriented situations than those from Germany and the United States.

Existing research has also indicated that culturally matched ad appeals are more likely to be persuasive than are mismatched appeals. U.S. respondents found ads emphasizing individualistic benefits to be more persuasive, and found ads emphasizing family or in-group benefits to be less persuasive, than did Korean respondents. Also, Chinese participants responded more favorably to collective ad appeals (e.g., "share the moments of happiness") than to individualistic appeals (e.g., "the joy of self-expression"), whereas the reverse was true for U.S. participants. In addition, advertisements in Mexico that depict values that are consistent with the local cultural norms and roles (e.g., familial norms and roles in Mexican culture) elicited more favorable attitudes and purchase intentions than did ads that depict inconsistencies.

Product characteristics, such as whether a product is personal or shared (i.e., privately used vs publicly visible), have been found to moderate the effect of culture on the persuasiveness of ad appeals. Only for shared or socially visible products—those purchased or used with in-group members—did strong cultural differences emerge between Americans and Koreans in Han and Shavitt's 1994 study (Fig. 1) as well as between Americans and Chinese in other research.

2.1.2. Horizontal and Vertical Cultural Orientations

Within the individualism–collectivism framework, a distinction between horizontal and vertical societies has recently been introduced. Members of horizontal societies, such as Denmark and Sweden, value equality and view the self as having the same status as others in society. In contrast, members of vertical societies view the self as differing from others along a hierarchy, accepting inequality and believing that rank has its privileges. In comparison with the individualism–collectivism dimension, the horizontal–vertical dimension is less explored. However, recent studies have begun to examine the implications of the horizontal–vertical cultural dimension. In 2002, Nelson and Shavitt found that members of a relatively vertical society (the
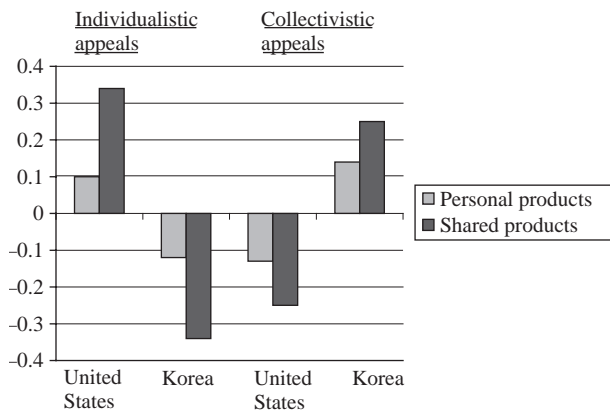
FIGURE 1 Persuasiveness of individualistic and collectivistic appeals in the United States and Korea. Bars represent mean standardized scores on an attitude index (the average of standardized scores across three evaluative measures). Adapted from Han and Shavitt (1994), *J. Exp. Soc. Psychol., 30*, 326–350, Copyright 1994, with permission from Elsevier.

United States) are more likely to endorse achievement values than are members of a horizontal society (Denmark). Moreover, Shavitt and colleagues' 2002 research suggested that vertical collectivism in the United States is associated with a tendency to be persuaded by ad appeals to status and hierarchy.

In 2000, Gurhan-Canli and Maheswaran demonstrated that Japanese consumers tended to favor products from their own country over foreign products, whereas American consumers tended to favor high-quality products regardless of the countries of origin of these products. Mediational analyses indicated that the vertical dimension of individualism and collectivism accounted for the country-of-origin effects in Japan. This is consistent with the cultural orientation in Japan—vertical collectivism—in that Japanese tend to believe in and value the superiority of the in-group over others.

In terms of advertising content, studies examining humor executions in ads have indicated that relationships between the central characters in ads where humor was intended were more often unequal in high power distance cultures (or vertical cultures such as Korea and Thailand) than in low power distance cultures (or horizontal cultures such as Germany); in the latter, these relationships were more often equal. Such unequal relationships in the ads are believed to reflect the hierarchical interpersonal relationships that are more likely to exist in high power distance (or vertical) cultures.

2.2. Self-Construal: Cultural Differences at the Individual Level

In 1991, Markus and Kitayama proposed that individualistic and collectivistic distinctions affect how individuals define or construct themselves in relation to others. According to these authors, the cultural norms of individualism require defining oneself as an individual whose behavior is organized and made meaningful primarily by reference to one's own thoughts, feelings, and actions. This independent self-construal is associated with describing oneself with reference to individual psychological traits such as "I am easygoing." On the other hand, the cultural norms of collectivism require individuals to see themselves as part of harmonious relationships and to recognize that their own behavior should seek benefits to their in-groups. This interdependent self-construal is associated with describing oneself with reference to social roles or memberships such as "I am in the gymnastics club." Recent research has indicated that self-construals account for and moderate cultural-level persuasion effects.

2.2.1. Chronic Self-Construal

Cross-cultural experiments have demonstrated that consumers with a more interdependent self-construal responded favorably to ads with a connected theme (e.g., interdependence, togetherness), whereas consumers with a more independent self-construal responded favorably to ads with a separate theme (e.g., independence, autonomy).

Self-construals have also been found to affect regulatory focus. Regulatory focus is built on the general hedonic notion that people approach pleasure and avoid pain. In 2001, Aaker and Lee found that, for people with an independent self-construal, promotion-focused information (e.g., drinking juice promotes energy or enjoyment of life) led to more positive attitudes toward a juice product. In contrast, for people with an interdependent self-construal, prevention-focused information (e.g., drinking juice helps to prevent disease) led to more positive attitudes.

2.2.2. Salient Self-Construal

Although an individual's cultural orientation may strongly determine the self-construal that is chronically accessible, self-construals may shift in response to other factors such as situational accessibility, motives, and reference points. The effect of temporarily accessible

self-construals on persuasion parallels that of chronic self-construals. For instance, regardless of whether self-construal is measured with the Self-Construal Scale or is situationally activated, studies have shown a consistent effect of independent self-construal versus interdependent self-construal on preference for either promotion-focused information or prevention-focused information.

The persuasion implications of salient self-construals have been further evidenced among bicultural consumers in the United States, that is, individuals who have been equally influenced by East Asian and Western cultural orientations. For example, Lau-Gesk's 2003 study found that an independent self-construal was temporarily activated when bicultural individuals were exposed to an individually focused appeal, and this in turn guided them to respond favorably to the advertised product. On the other hand, an interdependent self-construal was activated when bicultural individuals were exposed to an interpersonally focused appeal, and this in turn guided them to respond favorably to the advertised product. A frame-switching mechanism has been used to explain these findings in that bicultural individuals appear to be capable of shifting between individualistic and collectivistic interpretive frames in response to persuasive cues.

2.3. Culture and Cognitive Processing of Advertising Appeals

Just as psychological research on culture has focused extensively on cognitive process issues, recent research has demonstrated that a consumer's culture or cultural orientation influences the nature of information processing that accompanies a message.

Some studies along this line have focused on examining the effect of cultural orientation on persuasion by application of the Elaboration Likelihood Model/Heuristic–Systematic Model. Individualism and collectivism appear to be associated with different perceptions of cue diagnosticity in the processing of ads, and these in turn result in different patterns of attitudinal outcomes. Because of their individualistic tendencies, Americans are relatively uninterested in others' opinions; conversely, because of their collectivistic tendencies, people in Hong Kong tend to be more affected by the preferences of others. Consequently, among U.S. participants, product attribute information tends to be seen as more diagnostic than do consensus cues. In contrast, in 1997, Aaker and Maheswaran found that among Hong Kong participants, consensus cues are

treated as relatively diagnostic. Moreover, in 2000, Aaker and Sengupta's research demonstrated that individualistic versus collectivistic cultures are also associated with different strategies used in resolving incongruities between a source cue and product attribute information. In an individualistic cultural context, attribute information is perceived as more diagnostic than is a source cue and influences product evaluations. Hence, an attenuation strategy is used in information processing when incongruity between a source cue and attribute information occurs. In a collectivistic cultural context, where source information is relatively important and where there is less of a tendency to choose one type of information over another, the source cue and attribute information jointly influence evaluations. Hence, an additive strategy is applied when there is incongruity between a source cue and attribute information.

Culture also influences the role of affect in processing a message. Ego-focused (e.g., pride, happiness) versus other-focused (e.g., empathy, peacefulness) emotional appeals lead to more favorable attitudes among members of a collectivistic culture, whereas other-focused versus ego-focused emotional appeals lead to more favorable attitudes among members of an individualistic culture. This differs from the standard cultural congruity findings about persuasive effects. That is, relatively novel types of appeals, rather than culturally congruent appeals, elicit more elaboration. Because elaboration involves effortful thinking about the merits of a message as opposed to the use of simple cues, such elaborated thoughts tend to have a relatively large impact on evaluations.

2.4. Factors That Covary with Culture

Race, gender, and economic factors also affect ad content, ad processing, and attitude formation, although these effects are beyond the scope of this article. Very often, these variables will covary with culture, making studies of race, gender, and economic factors in conjunction with cultural ones particularly important.

The standard cultural congruity findings, suggesting that culturally matched ad appeals are more prevalent or are more likely to be persuasive, were reviewed earlier. However, a growing number of studies have indicated that the situation in rapidly transitioning economies may be more complex. For example, Westernized appeals, such as appeals to technology, individuality/

independence, and youth/modernity, have been found to be rather salient in Chinese and Japanese ads and to have been used frequently in current Taiwanese advertising agencies. These cultural incongruity findings could be driven by government policies regarding internal development and modernization, public exposure to Western media, and demographic and geographic contact zones. In addition, consumers in developing countries tend to respond favorably to Western products. For example, in one study of Indian consumers, brands perceived as having nonlocal (Western) countries of origin were favored over brands perceived as local. Moreover, this effect was found to be stronger for consumers who have a greater admiration for the lifestyle in economically developed countries. These cultural incongruity findings are meaningful because, as discussed by Zhang and Shavitt in 2003, they suggest the important role that advertising plays in reshaping cultural values in developing countries that are experiencing rapid economic growth. Rather than reflecting existing cultural values, advertising content in those countries promotes new aspirational values such as individuality and modernity; hence, these new values become acceptable and desirable among consumers.

3. CONCLUSIONS

As the advertising industry has become increasingly globalized, there has been increased research attention given to the role of culture in advertising. Progress in understanding this role has been made on a number of fronts. This article has reviewed empirical research on the effects of a number of cultural variables on advertising-related processes. Distinctions between individualistic and collectivistic cultures have been demonstrated to have important implications for advertising content, advertising persuasion, and consumers' cognitive responses. Parallel effects of culture and self-construal (independent vs interdependent) have been shown as well. In addition, studies on the distinction between horizontal and vertical cultural orientations within individualism–collectivism have begun to show advertising implications.

See Also the Following Articles

Industrial/Organizational Psychology across Cultures ■ Perception and Culture ■ Values and Culture

Further Reading

Aaker, J. L., & Lee, A. Y. (2001). "I" seek pleasure and "we" avoid pains: The role of self-regulatory goals in information processing and persuasion. *Journal of Consumer Research*, 28, 33–49.

Aaker, J. L., & Maheswaran, D. (1997). The effect of cultural orientation on persuasion. *Journal of Consumer Research*, 24, 315–328.

Aaker, J. L., & Sengupta, J. (2000). Additivity versus attenuation: The role of culture in the resolution of information incongruity. *Journal of Consumer Psychology*, 9(2), 67–82.

Gurhan-Canli, Z., & Maheswaran, D. (2000). Cultural variations in country of origin effects. *Journal of Marketing Research*, 37, 309–317.

Han, S., & Shavitt, S. (1994). Persuasion and culture: Advertising appeals in individualistic and collectivistic societies. *Journal of Experimental Social Psychology*, 30, 326–350.

Hofstede, G. (1980). *Culture's consequences.* Beverly Hills. CA: Sage.

Lau-Gesk, L. G. (2003). Activating culture through persuasion appeals: An examination of the bicultural consumer. *Journal of Consumer Psychology*, 13, 301–315.

Maheswaran, D., & Shavitt, S. (2000). Issues and new directions in global consumer psychology. *Journal of Consumer Psychology*, 9(2), 59–66.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implication for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253.

Nelson, M. R., & Shavitt, S. (2002). Horizontal and vertical individualism and achievement values: A multimethod examination of Denmark and the United States. *Journal of Cross-Cultural Psychology*, 33, 439–458.

Shavitt, S., Zhang, J., & Johnson, T. (2002). Horizontal and vertical orientations in cross-cultural consumer persuasion [abstract]. In S. Bronicarczy, & K. Nakamoto (Eds.), *Advances in consumer research* (Vol. 29, p. 47). Provo, UT: Association for Consumer Research.

Triandis, H. C. (1995). *Individualism and collectivism.* Westview.

Zhang, J., & Shavitt, S. (2003). Cultural values in advertisements to the Chinese X-Generation: Promoting modernity and individualism. *Journal of Advertising*, 32(1), 21–31.

Advertising Psychology

Patrick T. Vargas and Sukki Yoon

University of Illinois at Urbana–Champaign, Urbana, Illinois, USA

GLOSSARY

advertising A paid, mass-mediated attempt to inform and/or persuade.

AIDA model The acronym for Awareness, Interest, Desire, Action—a description of the stages by which advertising may have the desired effect on consumers.

attitude A psychological tendency to evaluate something (e.g., a person, an object, an idea) with some degree of favor or disfavor.

balance theory Theory that considers attitudinal relationships among a perceiver, an other, and a target; from the perceiver's perspective, balanced relationships are preferable to unbalanced relationships.

belief A piece of information about a person, an object, or an idea that an individual has in memory.

cognitive consistency A state when attitudes, beliefs, and behaviors are logically harmonious with one another.

cognitive dissonance An unpleasant state aroused by awareness of inconsistent cognitions.

Elaboration Likelihood Model and Heuristic-Systematic Model Dual-process theories of persuasion that posit both high-effort (central or systematic processing) and low-effort (peripheral or heuristic processing) routes to persuasion.

memory The mental ability to store and recall (intentionally or unintentionally) past experiences.

mere exposure effect Generation of favorable attitudes toward a given stimulus due to repeated (subliminal) exposure to that stimulus.

perception Recognition and interpretation of sensory stimuli.

persuasion The processes by which an individual's attitudes are changed.

priming A technique for increasing the accessibility, and likelihood of use, of some concept.

self-perception A theory suggesting that individuals infer their attitudes based on past behavior, particularly when they are uncertain of their attitudes.

subliminal advertising An attempt to advertise without the awareness of the message recipient.

Principles of psychological science have been applied to advertising practice since both fields were relatively young. The study of information processing, attitudes, and persuasion creates a foundation for advertising psychology because each is an important determinant of achieving advertising's main functions: to inform, persuade, and influence.

1. INTRODUCTION

The *Advertisers Handbook,* published in 1910 by the International Textbook Company, offered a wide range of advice to advertising professionals. According to the

handbook, an advertisement should not be too general, a headline should not be silly or deceptive, and an ad should be arranged logically, be concise, contain a proper amount of matter for the commodity being sold, and so forth. A few pages of the handbook were even devoted to the use of psychology in advertising: "Study of the goods or service to be sold is highly important, but no more important than the study of that wonderful subject, the human mind. The advertiser will do well, in all his work, to give special attention to psychological principles."

The handbook's advice has been heeded. Advertisers have used both psychological theory and method to gain a better understanding of consumers. The study of information processing has been particularly influential as advertisers have sought to gain consumers' attention and ensure that consumers interpret and understand their messages and (hopefully) retain some of the information. The attitude construct has been particularly important in advertising psychology because advertisements are generally intended to create positive attitudes toward some objects and because attitudes are an important determinant of behavior. Thus, an understanding of attitudes is doubly important for advertisers given that attitudes are both desired outcomes and causes of desired behavior. Therefore, the psychological study of attitude change, or persuasion, is another crucial part of advertising psychology. This article reviews important aspects of psychology that also play an important role in advertising.

2. THE PSYCHOLOGICAL STUDY OF ADVERTISING

2.1. Historical Perspectives

The scientific study of psychology is roughly 125 years old, with the first laboratory of psychology being founded in 1879 by Wilhelm Wundt in Leipzig, Germany. Modern advertising is only slightly older than scientific psychology, having appeared in the earliest newspapers circa 1850, reaching a circulation of approximately 1 million readers. The two fields came together in the earliest application of psychological principles to advertising practice in America in 1896. At the University of Minnesota, Harlow Gale conducted systematic studies of the influence of ad placement within a page on attention, the impact of necessary versus superfluous words in headlines, and how various colors used in ads might influence readers' attention.

Gale's research did not receive wide attention, but several years later another researcher, Walter Dill Scott,

wrote a series of articles on psychological aspects of advertising. Scott's research on advertising psychology tended to focus on the concept of suggestion. He believed that advertising was primarily a persuasive tool, rather than an informational device, and that advertising had its effect on consumers in a nearly hypnotic manner. People were thought to be highly susceptible to suggestion, under a wide variety of conditions, so long as the suggestion was the only thought available to them. According to Scott, advertising was most effective when it presented consumers with a specific direct command. "Think Different," a suggestion used by Apple Computer, is a good example of a direct command, but it would not necessarily be an effective advertisement because it does not tell consumers how to think differently. A more effective advertisement might be "Use Apple Computers."

2.2. Contemporary Perspectives

It is interesting to note that the concept of suggestion went out of style for some time among scientific psychologists. It is somewhat unflattering to think of people as automatons who unthinkingly follow whatever instruction is given to them; however, contemporary scientific psychologists have once again embraced the notion that a fair proportion of human behavior is due to influences of which people are entirely unaware. One intriguing line of contemporary research has shown that simply asking people whether they intend to purchase a new car sometime during the next year dramatically increases the chances that they actually do purchase a new car during that year—even if they respond "no" to the initial inquiry. Remarkable as these ideas are, advertising has relied on a number of basic, tried-and-true psychological principles that are covered in the remainder of this article.

One very general model of advertising effects, the AIDA (Attention, Interest, Desire, Action) model, has roots in Scott's writings. In 1913, Scott proposed a model of advertising effects that had three stages: attention, comprehension, and understanding. Over the course of several decades, Scott's model evolved into the AIDA model, which is still in use. According to this model, advertising, or promotions more generally, must first garner the attention of consumers and help them to develop beliefs about the product or service. Second, advertising should create interest or positive feelings about the product/service. Third, advertising or promotions should instill in consumers a desire for the product/service, thereby helping them to form purchase intentions. Finally, consumers must be

convinced to take action, that is, to purchase the product/service. This article covers all of these aspects of advertising psychology and more. It is important to note, however, that the AIDA model represents a strong version of advertising effects.

2.3. Strong and Weak Models of Advertising Effects

Advertising can strongly persuade consumers into immediate buying, or it can have a more subtle effect by reinforcing people's existing propensities to buy certain brands. The psychological processes underlying these two mechanisms also differ. The strong model focuses on consumers' immediate psychological or behavioral reactions where explicit advocacy and rationales of advertising messages are vital. The weak model emphasizes brand awareness where advertising is viewed as a reminder of a brand or source of information. For example, according to the strong model, Mark may decide to go to the shopping mall immediately after viewing a television commercial that says, "40% off any purchase at Macy's." On the other hand, David may buy Coke instead of Pepsi because he has greater familiarity with the Coke brand name, although both brands were initially in his consideration set.

Whether the strong or weak theory provides a better explanation for advertising effects hinges on numerous factors. For example, strong effects are more likely to occur in "high-involvement" situations where advertising directly aims at changing individuals' attitudes. However, when consumers have a predetermined set of alternative brands or are in a "repeat buying" situation where purchase decisions tend to be habitual, weak reinforcement enhances long-term brand awareness, familiarity, salience, and brand associations. In any event, for advertising to have either strong or weak effects, it must first be perceived by prospective consumers.

3. PERCEPTION

The first issue is the process of getting the advertising "into" consumers. The sheer amount of stimuli around people is quite overwhelming. Consider all of the following stimuli to which people are constantly being exposed but are unconsciously "filtering out": the air rushing into their mouths or nostrils as they breathe, the sensation of tightness where their shoes are touching their feet, the color of the walls nearest

them, and so forth. When one considers all of these things, it is remarkable that any advertising makes it into people's consciousness at all. Nevertheless, if advertising is to reach consumers, they must first and foremost be exposed to it.

3.1. Exposure

Exposure is the first stage in perception, and it occurs when people achieve physical proximity to some stimulus. Whenever something activates their sensory receptors, it means they have been exposed to that something. A printed advertisement must activate rods and cones in people's eyes, and a radio advertisement must move the small bones that make up people's inner ears.

This activation of sensory receptors may occur at different threshold levels, with the lower, terminal, and difference thresholds being important to advertisers. The lower threshold is the minimum amount of stimulus that is necessary for people to be aware of a sensation. Images presented very briefly (i.e., a few milliseconds) fall below the lower threshold and are imperceptible to people. The terminal threshold is the point at which additional increases in stimulus intensity produce no awareness of the increase. The difference threshold is the smallest noticeable change in stimulus intensity. These threshold levels have aroused interest among advertisers due to their application to subliminal advertising.

During the late 1950s, Jim Vicary claimed to be able to influence people without them even knowing that they had been exposed to influence attempts. Vicary claimed to have increased soft drink and popcorn sales by subliminally presenting messages to movie viewers. Years later, Vicary confessed to fabricating his claims, but the notion that people could be subliminally influenced took hold somewhere in the popular imagination. Sporadically throughout the next several decades, the question of whether subliminal presentation of persuasive messages could influence consumer behavior was debated. It is now known that subliminal presentation of stimuli can influence attitudes and behavior, but not quite in the way that Vicary suggested.

In 1968, Robert Zajonc discovered that positive attitudes can be induced simply by repeatedly exposing people to a subliminally presented stimulus, a finding that he termed the "mere exposure" effect. Participants in these studies were repeatedly exposed to a set of unfamiliar stimuli (e.g., Chinese ideographs), although the exposures were so fast that participants were unaware of even having seen the stimuli. Yet, when asked to evaluate a variety of similar stimuli, some of

which they had been exposed to and others of which had not been presented to them, they evaluated the stimuli to which they had been previously exposed more positively than they evaluated similar stimuli that had not been presented to them subliminally. Of course, developing favorable attitudes toward something is a long way from actually getting people to buy something; however, it is now known that subliminal presentation of stimuli can also influence behavior.

Individuals can be "primed" by the subliminal (or even supraliminal under certain conditions) presentation of stimuli in essentially the same way that they prime a pump (i.e., by filling it with fluid so that it is ready for immediate use). Following the presentation of a primed concept, individuals are more likely to use that concept given an appropriate opportunity. For example, individuals primed with the concept of elderly were observed to walk slower than individuals primed with neutral concepts. However, priming is a relatively weak influence on behavior and simply could not make people get up in the middle of a movie to go fetch soft drinks and popcorn. There are far more effective ways in which to influence consumer behavior.

3.2. Attention

One of the enduring problems facing advertisers is attracting consumers' attention. "Zipping and zapping" behavior is extremely common; television viewers record programs and zip (fast-forward) through the commercials or zap (eliminate commercials during recording or change the channel when they appear) them altogether. Drivers change the radio station when advertisements come on, they look away when billboards annoy them, and so forth. Attention is simply the conscious allocation of cognitive resources to some stimulus, but it is hard to grab.

A number of strategies have been used by advertisers to attract consumers' attention. Theories of human motivation have been popular guidelines for advertisers. For example, Maslow's hierarchy of needs posits five levels of motives that are common to all people: physiological (e.g., food, sex), safety (e.g., protection from the elements), belongingness (e.g., acceptance by others), esteem (e.g., admiration of others), and self-actualization (e.g., fulfillment of potential). Advertisers have tried to attract attention by positioning their products to fulfill one or more of these needs. Exercise equipment and healthy food appeal to physiological needs, bicycle helmets and smoke alarms appeal to safety needs, mouthwash and acne creams appeal to belongingness needs, designer clothing and luxury automobiles appeal

to esteem needs, and computer software and military service have even been made to appeal to self-actualization needs. One might imagine that nearly any product can be positioned to appeal to nearly any need—all in the name of attracting attention.

Properties of the stimulus can also be manipulated to attract attention. Size, color, intensity, contrast, position, movement, novelty, and so forth all can be used to attract attention.

3.3. Interpretation

Interpretation is the way in which meaning is assigned to stimuli. People exposed to exactly the same stimulus may interpret it very differently. For example, they may perceive a fuzzy kitten as an adorable target for cuddling or as a loathsome beast that causes sneezing and hives. Individual, situational, and stimulus characteristics all can influence the interpretation of stimuli.

At the individual level, consumers' expectations, motives, and attitudes all can have a profound influence on they way in which they interpret advertisements. When clear cola was introduced, consumers expected it to taste different from the traditional caramel-colored colas; however, clear cola did not taste much different from brown colas, and many consumers were turned off by this. The violation of their expectations may have, at least in part, caused them to ultimately reject the clear cola. Motives and attitudes are particularly likely to have an influence on the interpretation of ambiguous stimuli. Advertisers who do not send clear messages may be opening themselves up to very different, and possibly unflattering, interpretations that reflect consumers' own motives and attitudes.

At a situational level, the context in which an advertisement appears can have an influence on how it is interpreted. Several years ago, a prominent newspaper put out a Sunday supplement that featured a lengthy article on starvation and drought in Africa, followed immediately by a lengthy pictorial on the coming season's high-fashion formal wear. The following week, there was a section in the same supplement that was devoted to outraged letters from readers and a formal apology by the newspaper. Rarely are pictorials on fashion the target of so much outrage, but the context made a huge difference.

At a stimulus level, interpretations vary a great deal as well. As noted previously, stimulus ambiguity opens doors for motives and attitudes to have their influence on interpretations. Other stimuli can have vastly different interpretations due to cultural differences.

Different colors and numbers have different culturally laden meanings in different countries. There used to be a brand of cold medicine called "666." This brand likely did not fare well among conservative Christian consumers; but it might be prized in cultures where the number six has a positive meaning.

4. MEMORY AND LEARNING

Memory is complex; we can remember toys that we wanted as children, yet we sometimes cannot remember what we did the weekend before last. The simple change of a single word (e.g., how fast was the car going when it "bumped" [or "smashed"] into the pedestrian?) can substantially alter our memories. Yet it has been estimated that over the course of a lifetime, the average human stores approximately 500 times the amount of information that is in a full set of encyclopedias. Given these seemingly strange contradictions regarding memory, what hope can advertisers have of getting their products, brands, and ideas into consumers' memories?

4.1. Encoding, Storage, and Retrieval

Memory involves three main processes: encoding (the process by which information is put into memory), storage (the process by which information is maintained in memory), and retrieval (the process by which information is recovered from memory).

Encoding may be visual, acoustic, or semantic. Visual encoding and acoustic encoding are self-explanatory; they are named for the sensory modality through which they operate. Semantic encoding refers to the general meaning of an event. For example, one might encode a television advertisement in terms of the visuals presented, the sounds that accompany it, or the general idea that there is a sale at the market.

Storage may be short term or long term. Short-term memory, or working memory, is of quite limited capacity and is used to hold information in consciousness for immediate use. Long-term memory is quite vast and can retain information for extremely long periods of time (e.g., some childhood memories last until death).

Retrieval also comes in different forms. Explicit memory is tapped by intentional recall or recognition of items or events. Implicit memory is the unintentional recollection and/or influence of prior experience on a current task. On implicit memory tests, respondents are unaware that memory is being accessed. Implicit memory is assessed in a variety of ways such as word fragment completion (words seen previously are more likely to be completed than are words not seen previously) and time savings for tasks that have been done before. Advertisers may be particularly interested in explicit memory because the ability to intentionally recall information serves as a good measure of advertising effectiveness.

4.2. Short- and Long-Term Memory

Short-term memory is of relatively little consequence to advertisers because it does not, by definition, have any "staying power." Short-term memory may be important for direct response advertising, where consumers are asked to call a phone number to order or request additional information. But there is little that an advertiser can do with short-term memory. The primary function of short-term memory is to allow people to perform mental work such as calculating a sums or remembering a telephone number until they can either dial it or write it down. The fairly recent trend of placing advertisements on shopping carts in supermarkets might be a good way for advertisers to keep their products in short-term memory while people are doing their grocery shopping. In general, however, for an advertisement to be effective, it must work its way into long-term memory.

One bright spot for advertisers hoping to etch their work in consumers' long-term memory is the fact that most theorists believe that long-term memory has nearly unlimited capacity. So, there is always room for more information. But what is the best way in which to achieve storage in long-term memory?

Some information is stored in long-term memory accidentally (i.e., unintentionally), but most information that makes it to long-term memory is encoded semantically. According to one model of memory, information is better remembered when it is thoroughly processed. To the extent that people elaborate on, and think deeply about, a particular piece of information, they are more likely to be able to recall it later. This notion of cognitive elaboration is important in the understanding of persuasion as well. Clearly, advertisers would do well to generate ads that cause people to think deeply or discuss extensively. According to another model of memory, retrieval is enhanced to the extent that learning and retrieval occur under similar conditions. This bodes poorly for advertisers given that the encoding and storage of advertising rarely occur under the same

conditions as the actual buying behavior. Advertisements tend to be absorbed most everywhere except the marketplace.

4.3. Processing Pictures Versus Text

Pictures enhance advertising effectiveness in several ways: They help to get the audience's attention, they provide information about the brand and product use, and they help to create a unique brand image. Because people examine the visual elements before the verbal elements, pictures in advertising are a useful tool as attention grabbers. For example, extensive use of sex appeal and celebrity endorsement in advertising can be understood in this context.

Images and text are processed differently. Pictorial information seems to be processed more holistically, whereas verbal information is processed more sequentially. Information conveyed by pictures is recalled and recognized more easily than is textual information. Therefore, the fact that more than two-thirds of print advertisements today have pictures covering more than half of the available space is not surprising. So, why are pictures more memorable than words?

According to Paivio's dual-coding model, different codes exist in memory for verbal material and pictorial material. Pictures tend to be remembered over verbal information because they activate both verbal codes and pictorial codes spontaneously. This ease of cognitive access leads to the facilitation of memory. In addition, ads with concrete, easily identifiable pictures (e.g., celebrities) or realistic pictures are more memorable than ads with abstract or unidentifiable pictures. Furthermore, consumers exhibit more favorable attitudes when they see identifiable objects (e.g., an Adidas logo, a picture of Michael Jordan) than when they simply view text (e.g., the combination of alphabetical letters P-U-M-A).

Memory effects of text vary depending on how difficult it is for perceivers to comprehend the text. When the text is complex, people engage in more elaborative processing, resulting in a greater number of associations in memory as well as improved memory. Finally, pictures and text interact. Congruency between pictures and ad copy enhances consumers' recall.

4.4. Repetition and Learning

Advertising repetition is generally known to enhance memory by strengthening memory traces because it increases redundancy and provides more encoding opportunities to process the message, leading to higher levels of brand name recall. Nevertheless, advertising repetition rarely works in such simple ways. For example, varying the interval between message repetitions affects memory of an advertising message. Although not without exceptions, memory for repeated material generally improves as the time between presentations of advertising material increases, particularly when there is a delay between the subsequent presentation of the stimulus and the memory test.

Another explanation of learning through repetition is derived from encoding variability theory, which predicts that presenting a series of ads containing slight variations of a theme enhances memory for the ad material. For example, according to this view, repeated presentation of a bottle of Absolut Vodka in different contexts helps consumers to retain its brand name in their memory.

Attention, recall, and brand awareness initially increase, then level off, and ultimately decline as the number of exposures increases. At least two explanations are available for this "wear-out" effect: inattention and active information processing. With increasing repetition, viewers no longer attend to the message, and this inattention causes forgetting. According to the active information processing perspective, an audience rehearses two kinds of thoughts: thoughts stimulated by the message reflecting message content (i.e., message-related thoughts) and other thoughts based on associations reflecting previous experiences (i.e., people's own thoughts). With the initial exposure to a message, people's thoughts tend to be message related, but at some level of repetition, people's thoughts stem mainly from associations that are only indirectly linked to the message. These thoughts are less positive toward the product than are message-related thoughts, primarily because the latter were selected to be highly positive.

4.5. Low- and High-Involvement Learning

Learning may occur either in a situation where consumers are highly motivated to process the advertising material or in a situation where consumers have little motivation to learn the material. For example, a woman reading an automobile magazine in a dentist's office is less involved than another woman reading the same magazine prior to purchasing a new car. Personal, product, and situational factors jointly affect the level of involvement. A different level of involvement is likely to follow, for example, depending on whether the individual perceives the advertised product as enhancing

self-image (i.e., personal factor) or entailing risk (i.e., product factor). And this whole perception of the product will again be influenced by whether the consumer views the product for personal use or views it as a gift (i.e., situational factor).

High involvement stimulates semantic processing, whereas low involvement is linked to sensory processing. Text-based information is better remembered when viewers are highly involved, whereas graphic-oriented information exerts a greater impact on viewers' memory when they have low involvement. The study of learning and memory is important for understanding how consumers obtain information about a product or service. Another primary goal of advertising is to persuade, so the next section considers the psychological study of attitudes and persuasion.

5. ATTITUDES

The attitude construct has been recognized as one of the most indispensable concepts in psychology. It is similarly crucial in the study of advertising. But what exactly is an attitude? For a psychologist or an advertiser, an attitude refers to an evaluation along a positive–negative continuum. For purposes of this article, an attitude is defined as a psychological tendency to evaluate an object with some degree of favor or disfavor. The target of an attitude is called the attitude object. When an individual says, "This pie is lovely" or "That car is no good at all," that person is expressing his or her attitudes. When the individual says, "This pie is fattening" or "That car has poor acceleration," the person is expressing his or her beliefs. Attitudes are different from beliefs, although the latter help to make up the former. Beliefs are units of information that an individual has. Beliefs may be facts or opinion, and they may be positive, negative, or neutral with regard to the target.

5.1. The Structure of Attitudes

Attitudes are generally not thought of as monolithic constructs; they are made up of conceptually and empirically distinct components. At a very basic level of analysis, attitudes have three important components: affective, behavioral, and cognitive. Affect refers to feelings and emotional components of attitudes. Behavior, of course, refers to behavior that an individual takes with regard to a target. Cognition refers to the beliefs or thoughts that an individual has about a target.

Affective, behavioral, and cognitive processes help to form attitudes. The mere exposure effect suggests one way in which positive affect may arise. Classical conditioning and operant conditioning are two additional ways in which affective processes influence people's attitudes. The continuous pairing of some stimulus and a reward (or punishment) creates positive (or negative) affect. In advertising, brands of clothing are nearly always paired with attractive models. The models are intended to create affectively positive feelings, and advertisers hope that people will be conditioned to like their brands of clothing.

Behavior also contributes to the formation of attitudes in that sometimes people infer their attitudes on the basis of their previous behavior. Self-perception theory posits that people infer their attitudes on the basis of their past behavior, particularly when they believe that their behavior has been freely chosen. For example, if someone points out that Jane always wears green, she may infer that she has some affinity for green. But if Jane always wears green because her school has a strict dress code requiring her to wear green, she is unlikely to infer that she has a favorable attitude toward green.

Cognition is another important antecedent of people's attitudes. A cognitively based learning process occurs when people acquire information about attitude objects. People may gain information through direct experience such as when a free trial product is sent in the mail or when a free sample is offered in a store. Or, people may gain information indirectly, for example, when a television commercial shows them the benefits of owning a particular make and model of automobile. People's beliefs about attitude objects have been proposed as a central determinant of attitudes. Indirect learning, or observational learning, is an important tool for advertisers. Consider any advertisement that shows a model using a product to benefit in some way. It is hoped that viewers will develop favorable attitudes toward the product by learning how others have benefited from the product.

So, attitudes generally have affective, behavioral, and cognitive components. However, it is not necessary for all attitudes to have all three components. Some attitudes may be based primarily on affective factors (e.g., attitudes toward tequila), and others may be based primarily on cognitive factors (e.g., most people probably feel mildly positive about photosynthesis due to the important functions performed by the process, but they probably do not have strong emotions about it).

One very influential model of the structure of attitudes is Martin Fishbein's expectancy-value model.

Fishbein proposed that attitudes are a multiplicative function of two things: (a) the beliefs that an individual holds about a particular attitude object and (b) the evaluation of each belief. According to the expectancy-value model, beliefs are represented as the subjective probability that the object has a particular attribute. The model can be expressed as a mathematical function:

$$A_o = \sum_{i=1}^{n} b_i e_i,$$

where A_o is the attitude toward the object, b_i is a belief about the object, and e_i is the evaluation of that belief. According to Fishbein, people's attitudes are typically based on five to nine salient beliefs. So, if a researcher wanted to know someone's attitude toward a particular brand of clothing, the researcher might ask that person to estimate the likelihood that a particular brand has a variety of attributes (e.g., fashionable, durable, well priced) and how positive or negative each of those attributes is. The researcher could then compute an estimate of the person's attitude by multiplying the pairs of scores and then summing the products.

The expectancy-value model also implies that persuasion is largely a function of message content. That is, favorable attitudes can be produced by making people believe that an object is very likely to have some desirable trait, by making people believe that some trait is very favorable, or by both. For example, an advertiser might endeavor to make people believe that its automobile is very reliable (i.e., influence the subjective probability of beliefs) or to make people believe that its automobile's ability to take turns at very high speeds is highly desirable (i.e., influence the evaluation of a particular attribute).

Although the expectancy-value model seems to be perfectly logical, it may seem surprising to suggest that all attitudes are based on a series of beliefs. Consider, for example, the mere exposure research discussed earlier. According to Zajonc, preferences need no inferences (i.e., people may like something without having any beliefs about it). Under some conditions, attitudes may be formed outside of people's conscious awareness, or attitudes may be directly retrieved from memory rather than "computed" based on a mental review of salient beliefs. However, it is generally accepted that highly elaborated attitudes are more influential than poorly elaborated attitudes. So far, the discussion of attitude structure has considered how different aspects of a single attitude relate to one another. Next, the discussion considers how different attitudes relate to one another.

One of the most enduring psychological principles developed during the 20th century is the simple notion that people have a desire for cognitive consistency. Cognitive consistency is the simple notion that beliefs and actions should be logically harmonious. If an individual believes that cats make good pets but hates her pet cat, she has beliefs that are inconsistent; if an individual believes that cats make good pets and she loves her cat, she has beliefs that are consistent. For most people, cognitive inconsistency is unpleasant, so they take steps to achieve consistency.

One consistency theory with many advertising-related applications is Heider's balance theory. Balance theory was initially applied to cognitive consistency between dyads (two units) and among triads (three units), but because most research has examined triads, this article focuses on this arrangement. The triad arrangement pertains to the attitudinal relationships among a perceiver (p), an other (o), and an attitude object (x). Consider the example where Cody has recently met an individual named Sam, and Cody likes Sam quite a bit. One afternoon, Cody learns that Sam loves to listen to country music. However, Cody cannot stand country music. How does the fact that Sam loves country music make Cody feel? Probably not very good; the triad of Cody, Sam, and country music is not balanced. However, if Cody loved country music, Cody liked Sam, and Sam loved country music, all would be simpatico. These ideas illustrate the basic tenets of Heider's balance theory. As can be seen in Fig. 1, there are eight possible sets of relationships among the triads: four balanced and four imbalanced. One simple way in which to identify whether a triad is balanced or not is to calculate the product of the three relationships. If the product is positive, the triad is balanced; if the product is negative, the triad is imbalanced.

The efficacy of well-liked, or celebrity, endorsers may be explained at least in part by evoking balance theory. The viewers of the advertisement are expected to have a favorable attitude toward the endorser (e.g., Britney Spears), and the endorser is clearly portrayed as having a positive attitude toward the advertised product (e.g., cola). To maintain balance, viewers also should adopt a positive attitude toward the cola. Alternatively, viewers could decide to dislike the cola and change their attitude toward Spears so as to maintain a balanced triad. Balance theory also helps to explain one way in which consumer trends migrate. People who become friends with one another often adopt attitudes similar to their friends'
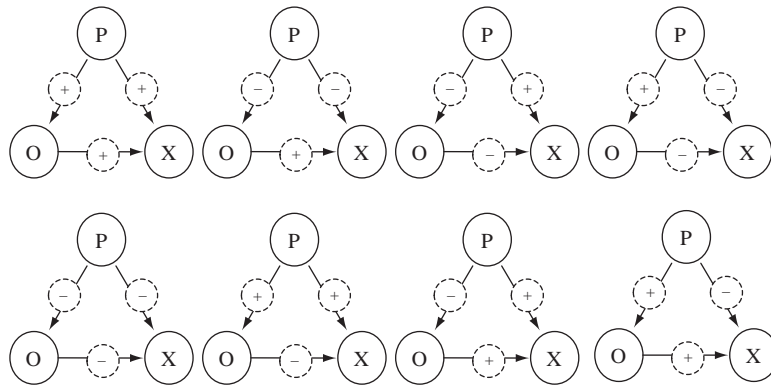
FIGURE 1   Eight possible triads proposed by Heider's balance theory. The triads in the top row are balanced, and the triads in the bottom row are imbalanced.

attitudes. A classic study by Theodore Newcomb illustrated this point with women who lived together at college; over time, the women's political attitudes became more and more similar.

Another theory that has roots in cognitive consistency, and has been very influential in advertising and consumer behavior, is cognitive dissonance theory. Leon Festinger proposed cognitive dissonance theory in 1957, and it spurred more research than perhaps any other social psychological theory. Cognitive dissonance has been defined as a feeling of discomfort that arises as a result of one's awareness of holding two or more inconsistent cognitions. Often, dissonance is aroused when one behaves in a manner that is inconsistent with his or her beliefs. For example, Greg may believe that Japanese cars are superior to cars made in America, but if he buys an American car, he will likely experience cognitive dissonance. Because cognitive dissonance is uncomfortable, people are motivated to reduce the feeling of dissonance by changing their behavior, trying to justify their behavior by changing their beliefs, or trying to justify their behavior by adding new beliefs. Having purchased an American car, Greg might try to reduce cognitive dissonance by investing in Japanese auto manufacturers, by changing his belief in the superiority of Japanese cars, or by adding a new belief to help regain consistency, for example, "My car may be American, but many of the engine parts are from Japan."

When people make large-scale purchases, they often experience what is known as postdecisional dissonance. Large expenditures may arouse dissonance because they are inconsistent with the need to save money or make other purchases. Furthermore, making a purchase decision necessarily means giving up some attractive features on the unchosen alternatives (e.g., buying a Sony means

not buying a Samsung). In a decision-making context, dissonance may be reduced by revoking the decision, by bolstering the attractiveness of the chosen alternative or undermining the attractiveness of the unchosen alternative, or by minimizing the differences between or among the alternatives. Another important role that advertising can play is in helping to reduce postdecisional dissonance. Advertising can help to reduce the feeling of discomfort that follows a major purchase by changing beliefs (e.g., "The new MP3 player has state-of-the-art technology") or by adding new beliefs (e.g., "The new MP3 player will make you the envy of your friends") that enable buyers to feel good about their recent major purchases.

5.2. Functions of Attitudes

Why do people have attitudes? Consider, for a moment, what life would be like without attitudes. For example, Tracy might sit down to dinner one day to find a plate of lima beans. She might eat the beans and discover that they taste horrible. The following week, Tracy is once again served lima beans and remembers them from the last time, but she has no attitude toward them. She eats them again and finds out, once again, that they taste quite horrible. For someone without attitudes, the discovery that lima beans are horrible could occur hundreds of separate times over a lifetime. At a most basic level, attitudes help people to navigate their world; they help them to know how to respond to things. Attitudes allow people to approach rewards and avoid punishments. Beyond this basic "object appraisal" function, attitudes have long been thought to serve a number of important functions.

More than 40 years ago, Daniel Katz described four functions of attitudes: ego-defensive, value-expressive,

knowledge, and utilitarian. Some attitudes have an ego-defensive function in that they help protect people from unflattering truths about themselves. People may bolster their own egos by holding negative attitudes about other groups (e.g., Hispanics, homosexuals). The value-expressive function occurs when attitudes allow people to express important values about themselves. For example, people may express a positive attitude toward recycling, suggesting that environmentalism is an important value. The knowledge function of attitudes allows people to better understand the world around them. For example, if a person dislikes politicians, it is easy to understand why politicians always seem to be giving themselves pay raises when the economy is particularly weak. Finally, attitudes may have a utilitarian function. Tracy's attitude toward lima beans helps her to know whether to approach them (because they are delicious) or avoid them (because they taste horrible). As suggested by all of these functions, attitudes also help to guide people's behavior.

5.3. Attitude–Behavior Relations

Indeed, the study of attitudes was at least partly initiated because attitudes seemed to be a logical predictor of behavior. As noted earlier in the article, behavior is an important form of attitudinal expression. For example, if Tara has a positive attitude toward a brand of pants, it follows that she will buy the pants; she will behave in a way that is consistent with her attitudes. Advertisers aim to create positive attitudes toward objects in the hope that consumers will purchase those objects. Unfortunately, the study of attitude–behavior relations has not been quite so simple. The following is only a very brief review of attitude–behavior relations.

A 1969 review of the literature on attitude–behavior relations found that attitudes and behaviors were modestly correlated with one another at best. This lack of attitude–behavior correspondence caused some researchers to suggest abandoning the attitude concept altogether. Fortunately, others rejected this suggestion and worked to better understand attitude–behavior relations. It is now known that attitudes reliably predict behavior under certain conditions. Attitudes and behaviors correlate when the attitude measures and behaviors correspond with regard to their level of specificity. If one wants to reliably predict a specific action, one should assess attitudes toward performing that action, with regard to a particular target, in a given context, and at a specific time. Or, one could enhance attitude–behavior correspondence by broadening the scope of

the behaviors. Knowing someone's attitude toward a particular attitude object (e.g., religion) is not necessarily going to predict whether that person attends church on a given Sunday, but it should reliably predict a variety of religious behaviors over the course of time (e.g., attending religious services over a period of weeks, having a religious text at home, wearing a symbol of religious faith).

Another way in which to enhance the attitude–behavior relationship is to provide people with direct experience with attitude objects. Indirect learning is not as powerful as direct learning, so an advertisement showing people enjoying a frosty beverage is not as persuasive as having people enjoy the frosty beverage in person. Behavioral prediction is enhanced when one also considers the influences of social norms and perceived control on behavior. Finally, attitudes are more predictive of behavior when attitudes are strong or readily accessible.

Attitude strength has been defined and measured in a variety of ways, but it is generally accepted that strong attitudes are resistant to change, persistent over time, and predictive of behavior. Clearly, these are the types of attitudes that advertisers want to inculcate. But given that there are multiple definitions of attitude strength, what is the best way in which to identify and produce strong attitudes? In considering various measures of attitude strength, it may be helpful to think in terms of whether the measures are operative or meta-attitudinal. An operative measure of attitude strength is one that reflects the operation of the attitude (e.g., the accessibility of the attitude can be measured by the speed with which people can provide evaluations of objects), whereas a meta-attitudinal measure of attitude strength is one that requires people to provide a subjective self-report of their own attitudes (e.g., the confidence with which an attitude is held cannot be directly measured but rather must be reported by an individual). One way in which attitudes can be strengthened is through cognitive elaboration. To the extent that people intentionally and carefully think about, and elaborate on, their attitudes, they are engaged in cognitive elaboration.

6. PERSUASION

The study of attitude change has existed since Aristotle's time. However, empirical research on persuasion and attitude change is a much more recent phenomenon. It was not until the mid-20th century that the study of attitude change developed into a thoroughly systematic process.

6.1. Message Learning Approach

Carl Hovland and colleagues, working at Yale University during the 1950s, sought to study persuasion by considering the question, "Who says what to whom with what effect?" That is, they were interested in studying the effects of different variables in the persuasion process. "Who" refers to the source of the persuasive communication, "what" refers to the message that is presented, and "to whom" refers to characteristics of the message recipient. This approach to the study of persuasion was an information-processing, or message-learning, paradigm. According to the message-learning paradigm, persuasive communications could have an effect only to the extent that they commanded attention and were comprehensible. Furthermore, message recipients had to yield to the persuasive communications and retain the information presented in the persuasive communications. If these conditions were met, beliefs, attitudes, and behaviors were liable to change. The following considers some of the critical variables studied by the Yale group.

Holding message, recipient, and channel characteristics constant, how do different characteristics of the source influence persuasion? Hovland and colleagues found that communicator credibility had an effect on persuasion, such that credible sources were more persuasive than noncredible sources. But what exactly is a credible source? Hovland and colleagues examined three characteristics of credible sources: expertise, trustworthiness, and the source's intent to persuade. An expert source (e.g., a Nobel prize-winning scientist) is more persuasive than a nonexpert source (e.g., the director of a local YMCA). A trustworthy source (e.g., a news anchor) is more persuasive than a nontrustworthy source (e.g., a member of the Liars Club). A source that is known to have a persuasive intent (e.g., an advertisement) is often less persuasive than one delivering the same message but with no persuasive intent (e.g., a friend). Forewarning people of a communicator's persuasive intent seems to instigate mental counterarguing in the audience. It is also known that physically attractive sources are generally more persuasive than unattractive ones and that similar communicators are usually more persuasive than dissimilar ones. Finally, powerful communicators (i.e., communicators who can administer punishments or rewards to the message recipients) tend to be more persuasive than powerless communicators.

Holding source, recipient, and channel characteristics constant, how do different characteristics of the message influence persuasion? The comprehensibility of a message is (obviously) an important determinant of persuasion; if people cannot understand the message, it is unlikely that they will be persuaded by it. The number of arguments in a persuasive communication also matters; more arguments are generally better than fewer arguments. Of course, there is a limit to the number of arguments one can present before message recipients become annoyed and lose interest. Presenting a few strong convincing arguments is better than presenting dozens of weak specious arguments. Messages that arouse fear (e.g., "If you do not brush your teeth, you will end up ugly, toothless, and diseased") can also be persuasive if certain conditions are met. An effective fear appeal must (a) convince the recipients that dire consequences are possible, (b) convince the recipients that the dire consequences will occur if instructions are not followed, and (c) provide strong assurance that the recommended course of action will prevent the dire consequences. One- and two-sided messages are differentially persuasive for different audiences; two-sided messages are generally more effective among knowledgeable audiences, whereas one-sided messages are more effective among less knowledgeable audiences. Two-sided messages can be more effective in general so long as the opposing arguments are effectively countered in the message.

Recipient characteristics also influence the efficacy of persuasive communications. An audience that is highly motivated and able to process information is more apt to be persuaded (provided that strong, rather than weak, arguments are presented) than is an audience that is distracted or apathetic. People who are of lower intelligence are generally easier to persuade than are people of high intelligence, people with moderate levels of self-esteem are generally easier to persuade than are people with either low or high self-esteem, and younger people are more susceptible to persuasive communications than are older people. Some personality traits also have important implications for persuasion. Self-monitoring is a characteristic that varies in the population. High self-monitors are particularly sensitive to situational cues and adjust their behavior accordingly, whereas low self-monitors are guided more by internal cues and tend to behave similarly across various situations. High self-monitors tend to be susceptible to persuasive communications that have image-based appeals, whereas low self-monitors tend to be more susceptible to value or quality-based appeals.

6.2. Dual-Process Theories

The message-learning approach to persuasion obtained a great deal of information about the influence of different

variables on persuasion. However, there were also a lot of apparently contradictory findings. One study might report that attractive sources had a large impact on persuasion, whereas another study would report no impact at all. During the late 1970s, a pair of integrative frameworks for understanding persuasion emerged. The Elaboration Likelihood Model (ELM) and the Heuristic–Systematic Model (HSM) are both dual-process models of persuasion, emphasizing that different variables can have different effects on persuasion. These models are fairly similar to one another and can be used to explain similar findings and make similar predictions, but the language used in describing the two models is different. The focus here is on the ELM because it has spurred the most research in advertising and marketing.

The ELM is based on the proposition that there is a continuum of elaboration likelihood. At one end of the continuum, the amount of cognitive effort that is used to scrutinize persuasive communications is negligible; information processing at this end of the continuum is minimal. This is known as the peripheral route to persuasion. At the other end of the continuum, there is a great deal of cognitive elaboration; people are highly motivated to carefully process information pertaining to the persuasive communications. This is known as the central route to persuasion. One might say that there are two routes to persuasion: central and peripheral. Which route is taken is dependent on message recipients' motivation and ability to process information. According to the ELM, people who are both motivated and able to process persuasive communications will take the central route, whereas people who are lacking either motivation or ability to process information will take the central route.

The central and peripheral routes are metaphors for the amount of cognitive elaboration in which people engage when they are faced with persuasive communications. Central route processing is particularly likely to occur when the persuasive communications are personally relevant to the recipient. If an individual is in the market for a new automobile, automobile advertisements would likely be personally relevant and, therefore, more highly scrutinized. If a person happens to be a home theater enthusiast, he or she would be more motivated to attend to advertisements for televisions and stereo speakers. Of course, all of the motivation in the world cannot cause greater cognitive elaboration if the message recipients do not also have the ability to process the information carefully. Distractions and other forms of cognitive business would prevent cognitive elaboration.

Persuasion through the central route occurs when the arguments presented are strong and compelling, and attitudes formed through the central route are persistent over time and resistant to change. Persuasion through the peripheral route occurs when there are compelling peripheral cues present (e.g., a long list of arguments, an attractive speaker, a credible source), and attitudes formed through the peripheral route are more temporary and subject to further change. The ELM and HSM can account for a wide variety of persuasion phenomena and have proven to be very robust and useful theories.

The psychology of advertising has come a long way. We now have a greater understanding than ever before of psychological processes crucial to advertising. From attracting attention to understanding how people are persuaded, the advancement of psychological science and that of advertising practice have much to learn from one another. This article has but scratched the surface of advertising psychology. New theories and applications are emerging at a rate that is unparalleled in the history of either advertising or psychology.

See Also the Following Articles

Attitudes ■ Environmental Versus Individual Risk Taking: Perception, Decision, Behavior ■ Learning ■ Perception and Culture

Further Reading

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley.

O'Guinn, T. C., Allen, C. T., & Semenik, R. J. (2003). *Advertising and integrated brand promotion*. Mason, OH: Thomson South-Western.

Packard, V. (1957). *The hidden persuaders.* New York: D. McKay.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 4, pp. 323–390). New York: McGraw–Hill.

Scott, L. M., & Rajeev, B. (Eds.). (2003). *Persuasive imagery: A consumer response perspective*. Mahwah, NJ: Lawrence Erlbaum.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.

Affirmative Action

David A. Kravitz

George Mason University, Fairfax, Virginia, USA

GLOSSARY

affirmative action Any measure, beyond a simple termination of discriminatory practice, adopted to correct for past or present discrimination or to prevent discrimination from recurring in the future.

affirmative action plan (AAP) A set of specific and results-oriented procedures designed to eliminate discrimination and enhance opportunities for members of an underrepresented group.

affirmative action plan (AAP) strength The weight given by an AAP to demographic status.

discounting principle In affirmative action, the tendency to derogate the competence of target group individuals who have been selected in the context of an affirmative action plan.

preferential treatment The provision of positive weight to minority status in an affirmative action plan; in weak preferential treatment (WPT), a member of the target group is preferred only if his or her qualifications are equivalent to those of a competing majority individual, whereas in strong preferential treatment (SPT), a member of the target group is preferred even if his or her qualifications are inferior to those of a competing majority individual.

principled conservatism theory The theory that opposition to affirmative action is not due to racial prejudice but rather is due to the inconsistency between affirmative action and conservative principles.

racioethnicity A term that refers to the combination of race and ethnicity that is used to distinguish among groups in affirmative action regulations.

racism The belief that a particular race is superior or inferior to other races.

social dominance orientation (SDO) A person's general approval of either dominance or egalitarian relationships among groups; individuals who are high in social dominance orientation support group-based inequality.

stigmatization In the context of affirmative action, the derogation of the competence or qualifications of an individual who is believed to have been selected on the basis of gender or racioethnicity.

utilization analysis An analysis that compares an organization's workforce (workforce analysis) with the availability of qualified individuals (availability analysis) to determine whether women or racioethnic minority groups are underutilized (i.e., whether there are fewer minorities or women in a particular job group than would reasonably be expected by their availability).

Affirmative action refers to any measure, beyond a simple termination of discriminatory practice, adopted to correct for past or present discrimination or to prevent discrimination from recurring in the future. This controversial social policy has inspired research in

a variety of disciplines. The present article summarizes research on (a) the economic effects of affirmative action on target groups and on organizations, (b) stigmatization of affirmative action target group members by others and by the individuals themselves, and (c) attitudes toward affirmative action.

1. WHAT IS AFFIRMATIVE ACTION?

1.1. Background

Humans have a propensity to distinguish between the in-group and the out-group, a process that often results in prejudice and discrimination toward individuals in the out-group. Consistent with this tendency, most countries have long histories of discrimination toward women and racioethnic minorities. During recent decades, many countries have passed antidiscrimination laws, and some have gone further to require that affirmative actions be taken to enhance opportunities for traditionally underrepresented groups. Although affirmative action may apply to university admissions, this article is restricted to affirmative action in the workplace. It summarizes the economic, behavioral, and psychological research on affirmative action that has been published in English. Most of this work has been conducted in the United States or Canada.

1.2. Definitions

What is affirmative action? That question can be answered conceptually, legally, and operationally.

1.2.1. Conceptual Definition

Many conceptual definitions have been offered. For example, the U.S. Civil Rights Commission has defined affirmative action as "a term that in a broad sense encompasses any measure, beyond a simple termination of discriminatory practice, adopted to correct for past or present discrimination or to prevent discrimination [from] recurring in the future."

1.2.2. Legal Definitions

The legal definitions of affirmative action vary from place to place and change over time. Critical dimensions on which the law varies include which groups are targeted, which actions are required and forbidden, what penalties can be imposed for violations, and

how the law is enforced. In the United States, federal affirmative action law is jointly determined by the U. S. Constitution (5th and 14th amendments), presidential executive orders (e.g., Executive Order 11246 [EO 11246]), legislation (e.g., Civil Rights Acts of 1964 and 1991), implementing regulations promulgated by federal agencies (e.g., Office of Federal Contract Compliance Programs [OFCCP], Equal Employment Opportunity Commission [EEOC]), and court decisions.

Organizations in the United States may be required to establish affirmative action plans (AAPs) for four reasons. First, courts may impose AAPs on organizations that have been convicted of discrimination. Second, to avoid adverse legal decisions, organizations may set up AAPs as required by court-approved consent decrees. Third, EO 11246 and other acts require affirmative action of federal agencies and of most organizations that do business with the government (i.e., federal contractors). Fourth, government "set-aside" programs at the federal, state, and local levels require that some proportion of contracting and procurement expenditures go to designated firms. In addition to these legal bases for affirmative action, some organizations may voluntarily take affirmative actions to diversify their workforces. All AAPs are constrained by the Constitution and by the Civil Rights Acts.

It is not possible to describe here the requirements and constraints of affirmative action in detail. However, because most AAPs in the United States are established by federal contractors as required by EO 11246, some of their key requirements and constraints are described here. First, the organization must not discriminate on the basis of race, color, religion, sex, or national origin. Second, the organization must complete an annual utilization analysis that compares its workforce, in terms of gender and racioethnicity, with the available qualified workforce. Third, if underutilization of women or a racioethnic minority group is revealed, the organization must generate a plan for eliminating the underutilization. Fourth, the plan should use targeted recruitment, outreach, and other nonpreferential actions that will facilitate target group success. Fifth, the plan must not involve the use of demographic status (i.e., racioethnicity, gender) in making selection, promotion, or termination decisions. This last point is important: Preferential treatment of women and racioethnic minorities is contrary to EO 11246. Preferences are legal only when they are specifically required by a court order or consent decree to reverse the effects of proven discrimination.

1.2.3. Operational Definitions

For current purposes, we must consider two types of operational definitions: organizational implementation of affirmative action and the descriptions of affirmative action used in behavioral research. Organizational AAPs are complex. They potentially include organizational structures, written reports, proscribed and prescribed processes, and more. Although "best practice" reports exist, the author has never seen a report summarizing the frequencies of typical practices. Importantly, although much of the public equates affirmative action with preferential treatment, the actual frequency of such illegal procedures is unknown. Operationalizations used in behavioral research tend to be brief, simplistic, and (often) illegally strong.

1.3. Public Beliefs

The American public construes affirmative action in many different ways. Some people focus on the role of affirmative action in decreasing discrimination against women and ethnic minorities. Other people believe it involves quotas or preferential treatment. Very few people think of affirmative action in terms of recruitment and outreach, and virtually nobody has a realistic understanding of the law. Beliefs about what the term "affirmative action" means vary with respondent characteristics, such as racioethnicity and political orientation, and predict support or opposition to affirmative action.

2. ECONOMIC EFFECTS OF AFFIRMATIVE ACTION

There has been a lively debate regarding the economic effects of affirmative action on the targeted groups and on organizations subject to affirmative action regulations. The implicit question is whether the benefits of affirmative action outweigh the costs. That question cannot be answered in the abstract because it depends both on the effects of affirmative action and on the values one assigns to organizational outcomes and to the well-being of the target groups. Although such values are a personal matter, it is possible to study the effects of affirmative action. Both conceptual and empirical studies have been published, and the conclusions are summarized in what follows. However, it is important to acknowledge some unavoidable methodological challenges. Causal conclusions are possible only when the experimenter can randomly assign

experimental units to treatment conditions while holding constant all other potentially relevant factors. That is impossible when dealing with a public policy such as affirmative action, where assignment to "conditions" is determined by the law and by the experimental units (organizations and target groups) themselves.

2.1. Target Groups

Many studies assess the employment status of White women and racioethnic minorities, vis-à-vis White men, in federal contractors versus noncontractors while controlling for relevant variables such as the type of industry. A positive effect of affirmative action is inferred if the relative outcomes of target groups are better in federal contractors than in noncontractors. This research indicates that affirmative action has had a positive effect on the employment status of White women and racioethnic minority groups, but the results are complex and some are tentative. At a minimum, the effects have varied across target groups, locations, time, and types of positions. The initial effect of affirmative action was substantial for African Americans and slightly positive for White women. The limited work on Hispanic Americans suggests that they were also helped. The clearest effects were on employment, especially in professional positions, although wages may also have been affected. However, this conclusion must be qualified by effects of region and time. The largest initial impact of affirmative action (and of antidiscrimination law in general) was observed in the South, where the extent of discrimination was most extreme and so there was the greatest room for improvement. In addition, changes were most substantial during the decade from 1965 to 1975. Broad progress stagnated during the 1980s, especially for African Americans, although there is some indication of continuing effects on the status of White women.

Other studies use micro-level employer data drawn from a variety of sources, including the organizations themselves. This work has found that target groups profit more from affirmative action than from passive equal opportunity practices. The effects of affirmative action vary across target groups, locations, industries, implementation procedures, and other factors.

A final indication of the effectiveness of affirmative action is the finding that growth in employment of White women and African Americans within organizations is predicted by organizational AAP goals. Although the relation is positive, the actual growth in employment of these groups is just a fraction of the

goals. This demonstrates that affirmative action goals do not function as strict quotas.

2.2. Organizations

The public debate surrounding affirmative action often centers on the presumed consequences of affirmative action policies for organizations. Critics claim that AAPs adversely affect performance by (a) forcing organizations to expend resources on implementation, (b) constraining organizational decisions, and (c) forcing organizations to hire underqualified individuals. Supporters of affirmative action argue that AAPs improve organizational performance by (a) motivating both White male and other employees and (b) providing organizations with the resources represented by a diverse workforce.

Theoretical models show that whether the net effect of affirmative action on organizations is positive or negative should depend on several factors, including the extent of preexisting discrimination, exactly how the AAP is implemented, and differences in human capital resources of majority and minority groups. Preexisting discrimination would hinder firm performance by limiting access to the resources represented by the victimized groups. If affirmative action decreased this discrimination, the effect on firm performance would be positive. On the other hand, if discrimination were already absent, the constraints on behavior posed by affirmative action could have a negative effect. Similarly, AAPs such as focused recruitment that serve to increase organizational access to human resources should have a positive effect on performance, whereas preferential selection should have a negative effect. Finally, the magnitude of the negative effect of preferential selection should increase with the size of the human capital gap between the majority and minority applicant groups.

As is clear from the preceding, there are logical arguments for both positive and negative effects of affirmative action on organizations. Thus, we turn to the empirical literature. The strongest research has included a wide variety of organizations and has controlled for industry, job level, and other influential variables. The paper qualifications (e.g., education) of minorities and White women appear to be somewhat lower than those of White males in firms with AAPs, particularly when affirmative action is used in hiring rather than when it is used only in recruiting. (The use of affirmative action in hiring does not necessarily imply preferential selection. Instead, it can involve actions such as the use of different screening mechanisms and the insistence that each

set of interviews must include one person who belongs to an underrepresented group. Many organizations reported that affirmative action was involved in hiring processes that resulted in the selection of White males.) Importantly, the minor decrements in educational qualifications are not associated with decrements in performance. Several explanations are offered for this apparent inconsistency, but all relate to the fact that firms that use affirmative action tend to have superior human resources practices. For example, virtually all AAPs include extensive recruitment, and this increases the overall quality of the labor pool. In addition, these organizations tend to use selection procedures that are unusually comprehensive and valid. Their use of comprehensive selection systems enables them to detect high-potential minority applicants whose competence is greater than their education would suggest. The null effect of affirmative action on organizational performance that has been observed in the multi-industry research has been replicated in the few studies that have assessed a single type of organization (e.g., police departments).

3. AFFIRMATIVE ACTION AND STIGMATIZATION

Concerns about stigmatization have been raised by judges rendering decisions in affirmative action cases, by Black and Hispanic intellectuals, and by conservative opponents of affirmative action. Empirical research has addressed several questions, of which three have received the lion's share of attention. Are members of affirmative action target groups stigmatized by others? Do members of affirmative action target groups expect to be stigmatized by others, and if so, what are their reactions? Do affirmative action beneficiaries engage in self-stigmatization? Work in this area draws on concepts from stigma theory and from attribution theory, particularly the principle of discounting.

According to the discounting principle, an individual's confidence regarding a possible cause of some outcome is affected by the presence or absence of plausible alternative causes. For example, if a woman is hired by a firm that has an AAP that targets women, she could have been selected either because she is highly qualified or to help the firm meet its affirmative action goal. In the absence of affirmative action, the only remaining explanation is her qualifications. Thus, a woman hired in the context of an AAP will be

evaluated less positively than a woman hired in the absence of affirmative action or less positively than a man. This reasoning applies both to observers and to the selected individual herself.

Note that the logic of discounting requires the assumption that affirmative action involves preferences. Also note that this reasoning assumes there are only two plausible reasons for an individual's selection—qualifications or affirmative action—and that a White male's selection must necessarily be based on his qualifications. This ignores factors such as traditional gender or racial discrimination, nepotism, personal friendship, similar backgrounds, and the myriad of other factors that are unrelated to merit but nonetheless affect employment decisions.

3.1. Stigmatization by Others

In the prototypical stigmatization experiment, the participant is given information about the background and performance of an employee who is either male or female and, if female, either is or is not an affirmative action hire. The affirmative action condition nearly always involves absolute preferences. The participant is asked to evaluate the employee on competence and other dimensions. Comparisons among the three conditions—male, female, affirmative action female employee—provide the key tests of stigmatization. In a second approach, the male participant believes he is working with a female leader on a one-way communication task, with the leadership role having been assigned on the basis of either test performance or gender. Again, the affirmative action condition involves explicit preferences. After successful or unsuccessful completion of the task, the participant is asked to evaluate his partner.

Consistent with the discounting principle, the affirmative action female is stigmatized. In the scenario research, she is evaluated less positively than either the male or the non-affirmative action female. In the communication research, she is evaluated less positively when selected on the basis of gender than when selected on the basis of test scores. Although the target is usually a White female, a few studies have included racioethnic minority targets. The stigmatization effect generalizes to these targets.

Having established the existence of stigmatization, subsequent research has concentrated on determining its limitations. Key conclusions include the following. First, stigmatization can be eliminated by providing incontrovertible information regarding the employee's competence or successful performance on the job.

Second, ambiguous information about employee success does not eliminate stigmatization. Third, stigmatization occurs when affirmative action is explicitly described as involving preferences and also when it is undefined. The latter finding is consistent with the common assumption that affirmative action involves preferences. Fourth, stigmatization does not occur when affirmative action is explicitly limited to the elimination of discrimination.

Much of this work has assessed the impact of respondent gender and racioethnicity and has sometimes found that stigmatization is most evident among respondents who are not members of the target group. With this exception, there is little research on the potential importance of other individual attributes, such as political ideology, prejudice, and belief in a just world, regarding the extent of stigmatization. In addition, there is little use of attribution theory other than the principle of discounting.

3.2. Target Group Members' Anticipation of Stigmatization

Conceptual and empirical work on self-fulfilling prophecies and stigmas suggests that the anticipation of stigmatization will affect the individual's affective, cognitive, and behavioral reactions. The belief that others consider one to be inferior will usually be distressing. It can stimulate either assumption-confirming or assumption-disconfirming behavior, depending on the situation and the individual. Indeed, concerns about stigmatization have motivated several Black and Hispanic intellectuals to speak out against affirmative action. However, little research directly addresses this issue in the context of affirmative action. The research that exists suggests that many White women and racioethnic minorities expect to be stigmatized by others even in the absence of affirmative action. They anticipate that affirmative action will increase the stigmatization, although they may express less concern about this possibility than do White males. The anticipation of stigmatization, with or without affirmative action, affects their feelings and behavior. One consequence may be a lessening of support for affirmative action. Nonetheless, they apparently believe that the known disadvantages of discrimination outweigh the possible disadvantages of stigmatization because they consistently support nonpreferential versions of affirmative action. A second possible effect could be an increase in stereotype threat. Stereotype threat refers to the fear that one may behave in such a way as to

confirm negative stereotypes of one's group. It is exacerbated by factors that draw attention to the stereotype and to the individual's membership in the devalued group. Although the presence of an AAP may be such a factor, this possibility has not yet been subjected to empirical investigation.

3.3. Self-Stigmatization by Target Group Members

Self-stigmatization refers to doubts regarding one's own competence, subsequent selection of a less demanding task rather than a more demanding task, decreased interest in the position in question, negative affect, and the like. There are two primary reasons to anticipate self-stigmatization as a result of preferential selection. The first relates to the discounting principle mentioned previously. If a woman knows that her selection is based on her gender rather than her qualifications, she may be less confident about her qualifications than she would be if merit were the only plausible reason for her selection. Preferential selection may also lead to self-stigmatization because receiving help can have self-threatening implications. The fact that one is helped suggests that one needed to be helped, and this implies incompetence.

Most research in this area has used the following research procedure. Male and female participants (typically White undergraduates) believe that they are about to partake in a one-way communication task. Assignment to the leader and follower roles is said to be based either on gender or on test performance. In reality, assignment to role is predetermined and the participant is assigned to the leader role. After role assignment, participants are asked about their leadership ability in general, their interest in serving as the leader in the communication task, and the like.

This research reveals consistent evidence of self-stigmatization by preferentially selected White undergraduate women. This effect is clearer for judgments of competence than for reports of motivation or task interest. However, when the AAP involves selection among comparable candidates, self-stigmatization is eliminated. Research on people of color is very sparse, but the work that exists has failed to observe self-stigmatization. This null effect likely stems from the belief by people of color that they are hurt more by discrimination than they are helped by affirmative action. Similarly, virtually all of this research has studied undergraduates; the limited work on employed adults has obtained no evidence of self-stigmatization. Several

studies failed to replicate the effect with White male respondents. This is consistent with the general lack of concern about self-stigmatization among White males, who have profited from traditional race and sex discrimination for centuries. These results suggest that initial self-confidence plays a critical role in reactions to preferential selection, and there is some empirical support for that perspective. Finally, psychological processes such as self-protective and egocentric biases would work against the development of self-stigmatization.

3.4. Summary

One observation applies to all of the work that has been discussed in this section: Negative effects are limited to situations in which affirmative action is operationalized as preferential treatment or is assumed to involve preferences. Preferential selection causes observers to question the competence of the selected individual and causes selected White women to doubt their own competence. These effects are eliminated or sharply reduced when affirmative action involves preferences among comparable candidates, and these effects are fully eliminated when affirmative action is restricted to the elimination of discrimination and opportunity enhancement efforts such as targeted recruitment.

4. ATTITUDES TOWARD AFFIRMATIVE ACTION

By far the largest body of research on affirmative action deals with the variables that predict attitudes. Support for affirmative action is frequently assessed in large surveys, and many dozens of questionnaire studies have been published.

Although many attitudinal predictors have been studied, this article is limited to the variables that have received substantial empirical attention. These predictors can be divided into three categories: structural aspects of the AAP, characteristics of the respondents, and psychological mediators. The structural factors include AAP strength, identity of the target group, and justification. The individual difference variables fall into two categories: demographic characteristics and opinion variables. Finally, the respondent's personal and collective self-interest and the perception of affirmative action fairness are presumed to mediate the relationship between the individual difference predictors and attitudes.

4.1. Structural Factors

The structural factors include details of the AAP. Real AAPs are complex. They include utilization analyses that compare the employer's workforce with the corresponding qualified labor market as well as organizational structures that assign responsibilities for maintaining equity. If target groups are underrepresented, the AAP includes goals for target group employment and a set of processes for implementation of the change. In contrast to this complexity, few research descriptions of AAPs exceed a single paragraph. In addition, much of this research includes AAPs that would be illegal if actually implemented.

4.1.1. Affirmative Action Plan Strength

The structural factor that has received the most attention is the manner in which the AAP attends to demographic status. This is typically referred to as AAP strength. The most common AAPs used in research, from weakest to strongest, include (a) the elimination of discrimination against the target group; (b) opportunity enhancement activities such as focused recruitment of the target group and the provision of training; (c) a tie-break procedure (often referred to as weak or soft preferential treatment) in which members of the target group are preferred if, and only if, their qualifications are equivalent to those of nontarget applicants; and (d) a preferential procedure (typically referred to as strong or hard preferential treatment) in which members of the target group are preferred even if their qualifications are inferior to those of nontarget applicants. Quotas would also qualify as strong preferential treatment. Although strong preferential treatment is normally illegal, this is the version most frequently used in behavioral research. Note that these four procedures confound the weighting of demographic status with the relevant phase(s) of the employment process.

Most research has found a negative effect of AAP strength on attitudes. The elimination of discrimination receives strong support. Opportunity enhancement procedures receive somewhat weaker support. Preferences are opposed, with most studies finding greater opposition to strong preferential treatment than to weak preferential treatment.

However, the preceding conclusions must be qualified. To a great extent, they are drawn from research on White respondents. When the effect of AAP strength on attitudes is studied within racioethnic minority groups, the negative monotonic effect observed among Whites is not always observed. Although all groups express opposition to strong preferences, racioethnic minorities make less of a distinction than do Whites among the weaker versions of affirmative action. Indeed, two studies on minority respondents have found greater support for recruitment than for the elimination of discrimination or any other AAP. Very few studies have included multiple racioethnic groups and AAPs of varying strength, so this conclusion is tentative.

4.1.2. Identity of Target Group

A few studies have manipulated the identity of the target group. There is a tendency for White respondents' attitudes to be less positive when the AAP targets minorities (especially African Americans) than when it targets White women or individuals with disabilities. Explanations for this effect could include racial prejudice (which would lower evaluations when the AAP targets racial minorities) and inclusion of the target group in the respondents' scope of justice (which would raise evaluations when the AAP targets White women or individuals with disabilities). However, there is such little work in this area that these conclusions must be tentative. One would expect the impact of target group status on attitudes to vary with demographic status of the respondents, but this interaction has received too little attention to permit any conclusion.

4.1.3. Justification

A number of studies have assessed the effects of explicit justification on attitudes. This work has found a modest positive effect of justifying the AAP on the basis of fairness (it is needed to remedy past discrimination) or economic value (it increases diversity, which has a positive impact on the organization's success). Justifying the AAP by stating that the target group is underrepresented appears to have a neutral or very small positive effect.

4.2. Individual Difference Variables

Dozens of studies have correlated affirmative action attitudes with a host of individual difference variables. The conclusions drawn in what follows are largely based on a meta-analysis of this research that was presented by David A. Harrison, David A. Kravitz, and Dalit Lev-Arey at the 2001 meeting of the Academy of Management.

4.2.1. Demographic Variables

Two demographic variables, racioethnicity and gender, have received the bulk of the attention, although there is also enough research on education to permit conclusions. Other variables such as age, income, and religion have occasionally been studied, but they have typically been included as controls rather than as variables of central interest. It is important to emphasize that demographic variables cannot serve an explanatory role. If one finds an effect of gender, for example, then one must ask why the effect occurred. The answer necessarily lies in the opinion variables that covary with gender and with the psychological mediators, as detailed in what follows.

4.2.1.1. Racioethnicity   One of the most consistent and powerful predictors of affirmative action attitudes is the respondent's racioethnicity. African Americans are the most supportive, Whites are the least supportive, and Hispanic Americans usually fall between the other two groups. The very limited work on Asian Americans suggests that their attitudes are usually more favorable toward affirmative action than those of Whites but are less favorable than those of other minorities. Despite these mean differences, it is important to note that disagreement exists within each of the groups. Many White Americans support affirmative action, and many African Americans oppose it. Also, as mentioned previously, attitudes vary with the interaction of racioethnicity by AAP strength. There is a large effect of racioethnicity in evaluations of strong preferential treatment, a weaker effect in evaluations of weak preferential treatment and opportunity enhancement plans, and no effect in evaluations of the elimination of discrimination.

Why is the effect of racioethnicity so consistent and substantial, at least in reactions to the stronger AAPs? One simple explanation is that attitudes are a function of self-interest, and affirmative action has different implications for the interests of the various groups. There is no doubt some truth to this explanation, but it is much too simple. For historical and cultural reasons, individuals in the various racioethnic groups see the world in different ways, and these differences affect their affirmative action attitudes. If we consider Blacks and Whites in the United States, for example, we find that African Americans are more likely than Whites to be politically liberal and to identify with the Democratic party. They are more likely than Whites to have experienced racial discrimination and to believe

that it is an ongoing problem. They are less likely than Whites to be racially prejudiced against Blacks and to believe that affirmative action involves preferences. Each of these differences is associated with attitudes toward affirmative action.

4.2.1.2. Gender   Numerous studies have assessed gender differences in attitudes toward affirmative action. In general, women are more supportive than men. However, this conclusion must be qualified. First, it is not clear whether it applies equally to all racial groups. It is only among Whites that gender determines target group membership, so self-interest considerations would suggest that the gender effect should be larger among Whites than among other racioethnic groups. At least one study has reported a reverse effect of gender among African Americans. Second, the author's reading of the literature suggests that the gender effect has shrunk over the past 20 years or so. However, the correlation between effect size and year of publication has not yet been reported, so this supposition is tentative.

4.2.1.3. Education   The third demographic variable that has received the most attention is education. Some of the most significant work on education has used it as a moderator of other predictors in the context of tests between competing theories. When the main effect of education has been assessed, the effect has varied with respondent race. The relation between education and support for affirmative action is usually positive among African Americans and is usually nonsignificant or negative among Whites. This latter result is anomalous because education is also associated with decreased racial prejudice and increased political liberalism, both of which predict support for affirmative action. This relationship between education and support for affirmative action among Whites merits closer investigation.

4.2.2. Opinion Variables

The two theoretical perspectives that have received the most research attention emphasize the importance of political ideology and prejudice. Distinguishing between these two approaches is complicated by the positive correlation between political conservatism and most measures of prejudice. A third theory posits that prejudice and political orientation serve as mediators between a desire for group dominance and affirmative action attitudes. Each of these three perspectives is described in what follows.

4.2.2.1. Prejudice Research indicates that prejudice is among the most powerful predictors of White Americans' attitudes toward affirmative action. However, the effect of racial prejudice varies with the measure used. The correlation is larger when racial prejudice is assessed with a measure of contemporary racism than when it is assessed with a measure of old-fashioned racism, and it is smaller yet when it is assessed as pure anti-Black affect. Because of a dearth of research, comparable distinctions among measures of sexism cannot be made.

The distinction between contemporary racism and old-fashioned racism requires some explanation. The core of old-fashioned racial prejudice, which prevailed among Whites until the past few decades, is the assumption of inherent inferiority of non-White racial groups. Blacks, in particular, were assumed to be inferior in intelligence and other characteristics. This belief was accompanied with support for racial segregation and for discrimination. Such beliefs have become relatively uncommon. Instead, many scholars argue, such old-fashioned racism has been replaced with contemporary forms of racism. The author uses the term "contemporary racism" to represent several different but related concepts, including ambivalent racism, aversive racism, laissez-faire racism, modern racism, racial resentment, subtle racism, and symbolic racism. These various conceptualizations differ in important ways, but most incorporate the following components. First, they assume that contemporary racists have underlying anti-Black feelings. Second, they assume that contemporary racists do not accept old-fashioned racist beliefs in inherent inferiority, segregation, and discrimination. Instead, contemporary racists believe that African Americans violate American values such as individualism and self-reliance. Third, contemporary racists acknowledge that racial discrimination existed in the past but believe that it is no longer a problem. Thus, they conclude that any economic disadvantages experienced by Blacks are due to their cultural inferiority. Finally, contemporary racists reject racism and any suggestion that their own beliefs are racist. One implication of this is that they feel most free to express their prejudice when some non-prejudicial reason for their action can be offered.

Given this distinction between old-fashioned racism and contemporary racism, one can understand why the latter predicts opposition to affirmative action more strongly than does the former. The anti-Black affect assumed by contemporary racism predisposes individuals to oppose affirmative action. The assistance that affirmative action offers to target group members decreases the importance of self-reliance. In addition, affirmative action is designed to eliminate the effects of discrimination that contemporary racists do not believe exists. Finally, because they can justify their opposition to affirmative action by arguing about the importance of self-reliance and the lack of ongoing discrimination, contemporary racists feel free to oppose it. On the other hand, old-fashioned racists who genuinely believe that Blacks are genetically inferior may conclude that African Americans need the assistance offered by affirmative action and so lessen their opposition. In addition, the ideological argument for individualism that is at the center of the contemporary approaches, and that is a core reason for opposition to affirmative action, is not central to old-fashioned racism. When both old-fashioned racism and symbolic racism are included as predictors of affirmative action attitudes, the latter is appreciably more important and typically explains incremental variance beyond the former.

One problem with the contemporary racism approach is that it overlaps conceptually and (in some cases) empirically with political conservatism. Thus, an individual who espouses a conservative ideology but genuinely lacks any prejudice toward Blacks would score relatively high on most measures of contemporary racism.

4.2.2.2. Political Orientation Beliefs associated with respondents' political orientation consistently predict attitudes toward affirmative action. Specifically, numerous studies have found that support for affirmative action is greater among Democrats than among Republicans and is greater among liberals than among conservatives.

From the beginning of the modern civil rights era during the 1950s, the major political parties in the United States have taken opposing positions on many racial issues. This divergence was solidified in the 1964 presidential race, when the Republican party nominated Barry Goldwater to run against President Lyndon Johnson. The civil rights movement played a central role in that election, and this established a pattern of Republican opposition to, and Democratic support for, government policies designed to assist racioethnic minorities. Indeed, in recent national elections, opposition to affirmative action, in particular, has been a formal or informal plank in the Republican party platform. Republican politicians typically refer to affirmative action as "preferential treatment" or "quotas" and argue

for its elimination. With this history, it is no surprise that party identification predicts public attitudes toward affirmative action.

Political conservatism is an even stronger predictor of affirmative action attitudes than is political party identification. Political conservatism includes several beliefs that engender opposition to affirmative action. First, conservatism posits that government should not intervene in the marketplace. Affirmative action is a federally mandated policy that constrains business actions. Second, conservatism is associated with support for self-reliance and individualism in both its normative and descriptive meanings. Normatively, individualism argues that individuals should take responsibility for their own success and should not rely on assistance from others (e.g., the government). Descriptively, individualism argues that success in the United States actually is determined by individual actions rather than structural forces. Affirmative action, with its underlying assumption that success is affected by structural factors and its targeted assistance to women and minorities, is contrary to individualism. A related point is that conservatives would support actions designed to ensure equality of opportunity but not those designed to ensure equality of outcomes. For these and associated reasons, conservatism is a strong predictor of opposition to affirmative action, particularly when it is construed as involving preferences.

The past decade has seen the development of a principled conservatism theory of affirmative action attitudes. The argument is that opposition to affirmative action is largely a function of race-neutral conservative principles rather than of racial prejudice. These scholars do not deny the importance of prejudice, but they argue that it is more important among the less educated than among the more educated, who are better able to behave in a manner consistent with their principles. One important point made by some principled conservatism scholars is that members of the public may be less than candid in their responses to survey questions regarding sensitive racial issues. When these researchers use both overt and covert measures of anger about affirmative action, they find the usual effect of political ideology on the overt measures, but this effect is greatly reduced or entirely eliminated on the covert measures. The effect of political orientation decreases because liberals express more anger on the covert measures than they do on the overt measures. A related point is that the effect of prejudice on affirmative action attitudes was stronger among liberals than among conservatives.

4.2.2.3. Social Dominance Theory   A number of scholars have developed models that emphasize the role of group conflict in motivating attitudes toward affirmative action. These approaches assume that societies are organized as hierarchies, with some social groups being dominant and others being subordinate. To maintain its position, the dominant group creates an ideology that legitimizes the status quo. In the current case, this ideology incorporates meritocracy, political conservatism, classical racism, and a belief in the cultural inferiority of racial minority groups. These beliefs serve as legitimizing myths; they legitimize the differences in status between the dominant and subordinate groups. They also legitimize opposition to affirmative action, which threatens the status hierarchy. This theory explains the correlation between prejudice and political conservatism, both of which serve as legitimizing myths, as well as the correlations of each with opposition to affirmative action.

Social dominance theory differs from some other conflict-based theories in its assumption that people vary in the extent to which they have internalized this perspective. This personality trait is called social dominance orientation. Supporting social dominance theory, there is a consistent correlation between social dominance orientation and affirmative action attitudes.

4.2.2.4. Comparison among Theories   There has been a rigorous debate within the affirmative action area among scholars who espouse the prejudice, principled conservatism, and social dominance perspectives. The resulting literature is quite complex and cannot be summarized briefly. In addition, much of it suffers from a significant limitation, that is, the equating of affirmative action with preferential treatment. As mentioned previously, attitudes toward affirmative action are strongly affected by AAP strength. More importantly, there is evidence that the importance of certain predictors varies with AAP strength. For example, research suggests that principled conservatism may be especially important in predicting reactions to strong forms of affirmative action that run counter to the basic principles of conservatism. On the other hand, racial prejudice seems to be most important in predicting opposition to weaker forms of affirmative action that are less inconsistent with conservative principles. To the author's knowledge, no research has assessed the relative importance of social dominance orientation across AAP strength.

The consistent correlation between measures of prejudice and opposition to affirmative action supports

both the prejudice and social dominance approaches. However, it is important to emphasize that the racism literature is complex and that the strength of the racism–attitude correlation varies with the conceptualization and measure of racism. The correlation is most substantial in studies that employ measures of contemporary racism that often overlap with measures of political conservatism. A limitation of the prejudice approach is that it does not explain attitudes of target group members.

Much of the research designed to competitively test these theories assesses the impact of potential moderators. For example, research has found that correlations between affirmative action attitudes and all three predictors increase with respondent education and political sophistication. This is fully consistent with the logic of social dominance theory, but the principled conservatism approach would predict a negative effect of education on the correlation between affirmative action attitudes and racism. Also contrary to the principled conservatism approach, the correlation between racism and political conservatism increases with education. However, these results are based on overt measures of racism and affirmative action attitudes. As mentioned previously, the principled conservatism scholars have demonstrated that the use of covert measures can profoundly affect results.

Some results supportive of social dominance theory were mentioned previously. In addition, research has shown that the significant correlation between racism and conservatism disappears when social dominance orientation is controlled. This result is consistent with the assumption that racism and conservatism are merely legitimizing myths that render the dominance hierarchy acceptable. Other supportive research has involved the use of structural equation modeling analyses of the social dominance and principled conservatism theories.

Finally, at least one study found that all three concepts—racism, conservatism, and social dominance orientation—contributed unique variance to the prediction of affirmative action attitudes. It is not now possible to say that one of these approaches is right and the others are wrong; it is likely that each is at least partly correct. It is also likely that the validity of these approaches varies with the operationalization of affirmative action. Unfortunately, little can be said about the moderating effect of AAP strength because, as noted previously, the vast majority of studies in the principled conservatism and social dominance areas have equated affirmative action with preferential treatment.

4.2.2.5. Beliefs About the Extent of Discrimination
More than 20 studies have correlated beliefs about the extent of target group discrimination with affirmative action attitudes. The result is clear: Support for affirmative action increases with the perception that the target group has suffered from discrimination. There are at least two reasons for this relationship. First, most people support equality of opportunity for all racial groups and actions designed to ensure such opportunity. Consistent with this reasoning, the association of perceived discrimination with support for affirmative action appears to be most substantial for the strongest AAPs. The second reason for this correlation is that perceptions of target group discrimination are associated with other predictors of affirmative action attitudes. For example, political liberals generally believe that discrimination is still a major problem, whereas conservatives tend to believe that discrimination is now uncommon. Along the same lines, social dominance theory would argue that disbelief in discrimination serves as a legitimizing myth.

4.2.3. Psychological Mediators

Finally, research has focused on the psychological mediators of affirmative action attitudes. Precisely how do AAP strength, racioethnicity, political orientation, and other variables affect attitudes? The two answers that have received the most attention are that they affect the respondent's perceptions regarding (a) the impact of the AAP on his or her interests and (b) the fairness of the AAP.

4.2.3.1. Self-Interest (Personal and Collective)
Assumptions regarding self-interest underlie the conflict theories, such as social dominance theory, mentioned previously. Self-interest has been assessed both indirectly and directly. An example of an indirect approach would be to use respondent age as an indicator of economic self-interest, assuming that younger workers are more strongly affected by affirmative action policies than are older workers. The direct approach asks the respondent to rate the likely impact of the AAP on his or her outcomes. Each of these approaches is flawed. The indirect measures suffer greatly from construct contamination and deficiency, whereas the direct measures suffer from common method variance and experimental demands. Research using indirect measures of self-interest finds that it is only weakly related to affirmative action attitudes, whereas work employing direct measures finds a strong relation.

Furthermore, research using direct measures finds that self-interest partly mediates the impact of other predictors such as gender and racioethnicity.

Self-interest has been studied at both the personal and collective levels. Personal self-interest refers to the impact of the AAP on the respondent's own outcomes. Collective self-interest refers to the impact of the AAP on the respondent's demographic group, which is typically defined in terms of racioethnicity and perhaps gender. The few studies that have included both predictors have found that they correlate similarly with affirmative action attitudes.

4.2.3.2. Perceived Fairness   Attitudes toward affirmative action are more closely related to perceived fairness of affirmative action than they are to any other variable. Indeed, research on evaluative reactions to affirmative action has often used measures of fairness rather than attitudes as the dependent variable. People report positive attitudes toward AAPs they consider fair and report negative attitudes toward AAPs they consider unfair.

Fairness perceptions mediate many of the effects described previously. Strong AAPs that provide rewards on the basis of demographic characteristics rather than merit are disliked because they are seen as unfair. That is a central argument of the principled conservatism perspective. Justifications affect attitudes in part because they affect the respondent's beliefs about the fairness of the AAP. In short, support for affirmative action is strongly predicted by beliefs about its fairness.

4.3. Complications and Limitations

As detailed previously, attitudes toward affirmative action are affected by actual or perceived details of the AAP, by the target group, and by justification for the plan. Attitudes are also associated with various demographic factors and opinion variables. Many of these effects are mediated by judgments of AAP fairness and self-interest implications. Most scholars would acknowledge the importance of all these predictors. There is disagreement, however, regarding their interrelationships and their relative importance. Unfortunately, most studies include relatively few predictors. In these cases, one cannot know how important the predictors would be in the context of additional variables. A related issue is that many of the most important predictors are individual difference variables that are intercorrelated and over which the experimenter has no control. This exacerbates the "missing variables" problem and also

means that causal conclusions cannot be drawn from the results. Of perhaps greater significance, not enough is known about how the predictors interact. We know that the relative importance of racioethnicity varies with AAP strength, but because of a dearth of research, it is impossible to draw strong conclusions about other interactions. This limitation is due in part to an unfortunate tendency for research to include only a single AAP.

The operationalizations of affirmative action have tended to be brief, simplistic, and (often) illegally strong. In many studies, affirmative action is not even defined. Instead, the stimulus is simply the term "affirmative action." In such situations, responses are determined by the participant's beliefs about what affirmative action entails, and these beliefs are rarely assessed. In addition, this work relies on respondent self-reports. Self-reports on controversial issues may be less than candid, especially if the research procedures do not guarantee anonymity. Furthermore, much of this research has used only White respondents. Even more disturbing, a few studies pool data across all racioethnic groups, and many studies pool results across minority groups. There is evidence that the minority groups differ in their attitudes. Of greater importance, it appears that minority groups may also differ in the factors that predict those attitudes.

Affirmative action is a complex social policy, and the determination of attitudes toward it is complex. Theories and research designs must be similarly complex if we are to obtain a fuller understanding of attitudes toward affirmative action. There is ample room for progress in this area.

5. CONCLUSIONS

One common thread weaves throughout the tapestry of research on affirmative action: Many negative effects are associated with the use of strong preferences but not with weak forms of affirmative action. Negative effects of affirmative action on organizational performance should be limited to situations in which the organization gives preference to underqualified target individuals. Stigmatization of target group individuals by others and by the individuals themselves occurs when affirmative action involves (or is believed to involve) preferences but not when preferences are known to be absent. Opposition to affirmative action and the resulting conflict exist primarily in response to preferences and quotas. The good news is that such

strong preferences are illegal, at least in the United States. The bad news is that a substantial segment of the population equates affirmative action with preferences. A major challenge for the future is to craft affirmative action procedures that are widely perceived to be fair and legitimate and that improve employment opportunities for traditionally underrepresented groups. Furthermore, the nonpreferential nature of these plans must be publicized, and the tendency to equate affirmative action with preferences should be eliminated in public discourse and academic research. Such actions should decrease social conflict while leading to a more equitable society.

Acknowledgments

See Also the Following Articles

Employment Discrimination ■ Prejudice and Discrimination

Further Reading

Bobocel, D. R., Son Hing, L. S., Davey, L. M., Stanley, D. J., & Zanna, M. P. (1998). Justice-based opposition to social policies: Is it genuine? *Journal of Personality and Social Psychology*, 75, 653–669.

Crosby, F. J. (2004). *Affirmative action is dead; long live affirmative action*. New Haven, CT: Yale University Press.

Crosby, F. J., & VanDeVeer, C. (2000). *Sex, race, and merit: Debating affirmative action in education and employment*. Ann Arbor: University of Michigan Press.

Donohue, J. J., III, & Heckman, J. (1991). Continuous versus episodic change: The impact of civil rights policy on the economic status of Blacks. *Journal of Economic Literature*, 29, 1603–1643.

Doverspike, D., Taylor, M. A., & Arthur, W., Jr. (2000). *Affirmative action: A psychological perspective*. Huntington, NY: Nova Science Publishers.

Federico, C. M., & Sidanius, J. (2002). Sophistication and the antecedents of Whites' racial policy attitudes: Racism, ideology, and affirmative action in America. *Public Opinion Quarterly*, 66, 145–176.

Heilman, M. E., Battle, W. S., Keller, C. E., & Lee, R. A. (1998). Type of affirmative action policy: A determinant of reactions to sex-based preferential selection? *Journal of Applied Psychology*, 83, 190–205.

Holzer, H. J., & Neumark, D. (2000). Assessing affirmative action. *Journal of Economic Literature*, 38, 483–568.

Holzer, H. J., & Neumark, D. (2000). What does affirmative action do? *Industrial and Labor Relations Review*, 53, 240–271.

Kravitz, D. A., Harrison, D. A., Turner, M. E., Levine, E. L., Chaves, W., Brannick, M. T., Denning, D. L., Russell, C. J., & Conard, M. A. (1997). *Affirmative action: A review of psychological and behavioral research*. Bowling Green, OH: Society for Industrial and Organizational Psychology.

Kravitz, D. A., & Klineberg, S. L. (2000). Reactions to two versions of affirmative action among Whites, Blacks, and Hispanics. *Journal of Applied Psychology*, 85, 597–611.

Major, B., Feinstein, J., & Crocker, J. (1994). Attributional ambiguity of affirmative action. *Basic and Applied Social Psychology*, 15, 113–141.

Rubio, P. F. (2001). *A history of affirmative action, 1619–2000*. Jackson: University Press of Mississippi.

Sears, D. O., Van Laar, C., Carrillo, M., & Kosterman, R. (1997). Is it really racism? The origins of White Americans' opposition to race-targeted policies. *Public Opinion Quarterly*, 61, 16–53.

Selmi, M. (1995). Testing for equality: Merit, efficiency, and the affirmative action debate. *UCLA Law Review*, 42, 1251–1314.

Sniderman, P. M., & Carmines, E. G. (1997). *Reaching beyond race*. Cambridge, MA: Harvard University Press.

Spann, G. A. (2000). *The law of affirmative action: Twenty-five years of Supreme Court decisions on race and remedies*. New York: New York University Press.

Age-Related Issues among Minority Populations

Edna E. Brown and James S. Jackson

University of Michigan, Ann Arbor, Michigan, USA

GLOSSARY

biopsychosocial model A framework for understanding human behavior, health, and well-being across the life span; it incorporates the biological, psychological, and social perspectives on the etiology, understanding, and treatment of illness and disease.

formal care Paid care received from private or government social service agencies or organizations.

informal care Unpaid care received from family, friends, neighbors, or religious organizations.

receiving care Help with impairments in activities of daily living (ADLs), such as eating, walking, using the toilet, bathing, and getting in and out of bed, or help with limitations in instrumental activities of daily living (IADLs), such as cooking, cleaning, and shopping.

reciprocity The exchanges of goods, services, and support of equal value among members of a social network.

sense of control The degree to which people believe that they can deliberately bring about desired outcomes and avoid undesirable ones in their environment.

In contrast to just a few years ago, more research on the psychological and social gerontology of racial and ethnic minority groups is being included within the general investigation of ethnicity and cultural factors in aging and human development. The existence of large national data sets and more powerful analytical techniques is increasing the quality and quantity of aging research on these groups in many areas of psychological interest. National longitudinal data collection efforts are improving the available data on the aging experience of minority Americans. Although better data are always needed, the improvement in a relatively few short years has been impressive. Similarly, the approach to research on the ethnic minority elderly has also seen a greater recognition of the heterogeneity among and within these groups. Research is now more focused on the role of culture, socioeconomic status, and gender as important markers of potential process differences within and among aging

groups of color, especially Black Americans. Recent ethnogerontology research is reversing historical trends, and generalizable high-quality findings are emerging concerning health, socioeconomic status, work and retirement, social support, well-being, and especially family patterns of caregiving and care receiving. This article reviews some implications of this work for the applied psychology of aging, with a special focus on issues related to perceptions of control, family caregiving, and care receiving among racial and ethnic minority populations.

1. INTRODUCTION

Research on racial and ethnic minority populations has propelled significant interest in the development of life course models and frameworks that assume an important role of race and ethnicity, culture, acculturation, and national origin in how various groups in the United States traverse the individual life course and age. The growth of racial and ethnic groups in the United States, and the relative deprivation of these groups in material, social, and health resources, demands greater attention if we are concerned about characterizing them appropriately and providing effective and adequate services as they age.

The growing internationalization of research on aging demands that we move beyond singular cultural perspectives. Fortunately, the United States is culturally heterogeneous, and comparative research on racial and ethnic groups yields excellent models relevant to understanding basic processes in aging in larger cross-national contexts. Important processes, such as stress and coping, may be more sharply defined in environmentally pressured groups (e.g., Blacks, Latinos, Native Americans). Racial and ethnic group boundaries may provide a personal and group coalescing of culture/environment and biological differences. For example, recent research supports degrees of assimilation and acculturation as important variables for understanding basic processes in mental health and mental disorder outcomes.

2. THE CHANGING NATURE OF AGING IN AMERICA

America has always been a nation of immigrants. Several factors will change the face of America in the future. The population will age dramatically. The racial and ethnic composition of the country will change significantly. Earlier waves of migration came from Europe, China, and Africa, whereas new waves are coming from Central and Latin America, the Middle East, and Southeastern Asia. Although the country is aging rapidly, the rate of aging is faster among the racial and ethnic minority "new" immigrant groups, especially Latinos, Asians, and (to a lesser extent) African Americans and other Black immigrant groups.

This changing racial and ethnic minority population will represent culturally unique groups of aging individuals. As culture has affected their lifetime experiences, it will also influence the nature of their aging experiences. Elements of the biopsychosocial model of aging are fundamentally affected by the cultural experiences of aging individuals, both cumulatively and contemporaneously. Immigrant groups in the United States have unique histories. For example, Cuban immigrants fled Castro and arrived in the United States feeling forcibly expelled from their homeland. They came to the United States, which was largely anti-Castro and sympathetic to their plight. Middle Easterners escaped dictators, war, and oppression, but (at least currently) they are aging in a considerably less sympathetic, if not hostile, United States. Asians (e.g., Koreans) also fled a homeland torn by war and remain embedded in communities with specific cultural values that guide their expectations about aging. The life span experiences of these individuals may include immigrating as children and spending 60 years in the United States, learning another language, living in urban versus rural communities, and being among a minority versus majority culture. Some immigrants left a higher standard of living, whereas others came to the United States to seek a higher standard of living. But particularly relevant to the aging experience is the fact that all immigrants have been influenced by their biological and physical characteristics, their family and friendship relationships, and the communities within which they live (i.e., their biopsychosocial development). Although this article focuses on minority racial and ethnic minority groups in the United States, these biopsychosocial influences are no less important for the majority group. In this case, recognizing the significant influences on racial and ethnic minority groups only serves to highlight the role of these factors for aging among all individuals.

3. A LIFE COURSE FRAMEWORK FOR AGING IN RACIAL AND ETHNIC GROUPS

As changing demographics require increased focus on older ages, lessons learned from the life span model of development are particularly germane. A life span framework is needed to explore how environmental stressors influence and interact with group and personal resources to both impede and facilitate the quality of life of successive cohorts of racial and ethnic minorities over the group life course and in the nature of their individual human development and aging experiences. Research on socioecological predictors of health in African American and Caucasian residents of a large midwestern city exemplifies the need for the life course approach. Findings have demonstrated relationships between socioecological factors (e.g., high crime rates, family dysfunction, high noise levels, social isolation) and negative health factors (e.g., hypertension) among minority residents. These factors can affect all members of the minority families and community, thereby possibly initiating poorer health among both younger and older individuals. It is the premise of a life course perspective that current and aging cohorts of ethnic minorities have been exposed to the conditions that will profoundly influence their social, psychological, and health statuses as they reach older ages during the years and decades to come.

Recent work on the incomplete ontogeny of human development and the recognition of the influence of culture in human development represents an important extension of the original work. In this latest consideration of life span developmental theory, the concepts of selection, optimization, and compensation are illustrative of the adaptive strategies that people use to maximize their competency. With the challenges of age, one has an increased need to select, optimize, and compensate to achieve designated goals. At the same time, it is noteworthy that culture is infused in every element of this process. What one selects, how one chooses to optimize, and what one considers appropriate forms of compensation all are culturally influenced if not culturally determined. As one ages, this influence accumulates and appears to have ever-increasing effects on the experiences of aging. The greater emphasis on culture in this life course modeling is very important in understanding the experiences of racial and ethnic minority groups. In fact, some have suggested that although culture becomes increasingly important

with age, its actual influence on development decreases with age. Thus, biological and physical development are said to dominate during childhood when culture has minimal impact. Within most Western cultures, this appears to be true. But of course, one might imagine a culture where one group is disfavored (e.g., girls) and, thus, is provided only very limited access (e.g., to nutritional resources), thereby limiting the biological or physical development that might otherwise dominate. Similarly, among older people, what one values and how one copes with biological or physical losses might also be influenced by culture.

The current authors believe that it is necessary to incorporate both the life span model and the role of culture, especially a view of culture nuanced by racial and ethnic differences, into life course biopsychosocial models, especially as they apply to aging.

4. RESOURCE DEPENDENCY: WORK, ASSET ACCUMULATION, AND RETIREMENT

4.1. Socioeconomic Status

Many minority elderly individuals continue to lag behind Whites in social and economic status. For example, indicators of income, education, and health status document the deprived position of Blacks relative to Whites. Some argued that this was a cohort effect, but the lingering poor relative position of Blacks refutes this; new cohorts of the Black elderly are not faring significantly better than priors ones. For example, there is a continuing disadvantage of middle-aged and younger Blacks (and increasingly Hispanic groups), relative to Whites, in housing, income, occupation, health, and education. Recent reports on the circumstances of Blacks across the entire life course continue to show the presence of relatively poorer circumstances, especially wealth, suggesting that new cohorts of elderly Blacks will continue to experience relative disadvantage in comparison with other groups.

On the other hand, even though approximately one-third of the Black elderly continue to live below the poverty level, today's Black elderly are better fed, better housed, and in better health than were those during earlier eras. Most of this improvement is attributable to government assistance programs, which is still the prime support of many black Americans and increasingly among other older age groups as well. Unfortunately,

because of histories of poor occupational opportunities, lack of wealth, and private retirement funds, a large proportion of Blacks are heavily dependent on these government programs. It is still unclear whether future cohorts of older Blacks will enjoy what may be the relative luxury of today's elderly. Recent diminution of federal programs, a change in the economy favoring job creation in "high-tech" and specialized educational intensive sectors of the economy (e.g., computers, communications), and simultaneous growth in low-paying service positions (e.g., fast-food restaurants) provide little room for today's Black adult and middle-aged cohorts, many of whom lacked basic educational opportunities during their formative years. It is now predictable that future cohorts of older Blacks may not be as well off as their White counterparts, and it is still unlikely that they will be as well off as today's cohort of Black elders, many of whom worked in union-intensive industries.

4.2. Health Morbidity and Mortality

At nearly every point from birth to death, African Americans and other minority groups (e.g., Native Americans) have poorer morbidity and mortality rates than do Whites. It is also well documented that there is increased longevity among Blacks (and some suggest that this is also true among Hispanics) who live to approximately 84 years of age. Many have suggested a possible selection bias favoring the long-term survival of particularly robust and hardy individuals (e.g., immigration) or differential rates of aging within Black, Asian, Hispanic, Native American, and White populations. Others have claimed that this crossover is only an artifact of faulty reporting and exaggerated age claims. The effect has been firmly established, although there is still no widely accepted explanation. The racial mortality crossover appears to be a real phenomenon—one that may involve some type of "survival of the fittest."

Recent research on the "oldest old" continues to document the heterogeneity of the social and psychological health of very old Blacks. This type of data provides strong support for a thesis that views the mortality crossover as involving the survival of hardier old Blacks and not a methodological artifact. Similarly, other work reveals some evidence for greater functional health among older elderly Blacks in comparison with Whites, although the effect seems highly dependent on educational status.

Recent research points to differences between ethnic minority groups and Whites in the nature of self-reported health. Early in the past decade, it was found that the largest differences were in the validity of the subjective interpretations of health state. These findings challenge traditional thinking and research regarding possible race differences in health. At this point, whether there are differences in the structure of health, the processes of health, or the influence of service use on experienced health problems remain open questions. What is clear is that changing health policies may have increasing negative effects on the ability of Black and many other minority elderly to receive adequate health care in the new century.

4.3. Psychological Well-Being

Research on psychological well-being has shown an increasing sophistication over the past few years. Structural factors, such as income and education, tend to show small but positive relationships to well-being. Some recent work suggests that psychological well-being may be strongly tied to family and health satisfaction. Some recent evidence also suggests that younger ethnic minority cohorts may be less satisfied than older cohorts at comparable periods during the life span. This is in sharp contrast to Whites, who have shown the opposite pattern. This lowered satisfaction and happiness in younger Blacks may be related to rising expectations and structural constraints that are likely to persist into older age.

4.4. Work and Retirement

Little empirical research had been devoted to the study of work and retirement among ethnic and minority elderly individuals. Some earlier work had speculated that the entire retirement process, viewed within a life span context, may be very different for Blacks and other disadvantaged ethnic minorities. For example, Blacks and other groups, such as Native Americans and some Latino groups, often have long histories of dead end jobs with poor benefits and bleak expectations, thereby lowering any advantages of retirement. Thus, inadequate income, poor housing, and uncertain futures may confront these older groups at retirement age. Faced with limited retirement resources, many older ethnic minorities may continue to work past customary retirement ages out of desperation. Some research indicates that these individuals are physically, psychologically, and socially worse off than their retired counterparts. As suggested earlier, even the relatively poor but stable government retirement

support (if these individuals are fortunate enough to qualify for such support) may, in contrast, be better than sporadic and poor jobs in the regular labor market. Thus, retirement may provide a small but secure government income, leading to increased psychological and social well-being.

4.5. Family and Social Support

Historically, research on minority family and social support networks has been based predominantly on anecdotal data. Two myths have dominated this area. The first is a view of older minorities being cared for by loving and extended family members and fictive kin. The other is a view of impoverished, lonely, older minority elderly individuals being abandoned by a disorganized and incompetent family system. National and other large social surveys indicate a reality somewhere in between, documenting the existence of extended family forms but also demonstrating that much of the assistance is reciprocal; that is, ethnic minority elderly individuals often provide help to younger family members and neighbors. Recent research supports the importance of community institutions, such as churches, as sources of physical and emotional support for older Blacks. This work points to the considerable obstacles faced by many ethnic minority Americans in providing services to physically disabled and/or demented (and mentally ill) relatives.

5. IMPORTANCE OF FORMAL AND INFORMAL CARE

5.1. Benefits of Support/Informal Care

Research has provided impressive evidence of the importance of social support throughout the life span, particularly to the elderly. Social support has salutatory effects on physical and mental health of older adults. Support reduces the negative effects of health problems, helps in coping with stressful life events, and buffers the impact of psychological distress. A lack of social support from family and friends is associated with poor psychological well-being (e.g., depression). The most important social support usually involves long-lasting, significant, and close relationships with a network of family, friends, and church members. The quality, rather than the quantity, of the support relationship has been shown to contribute most

significantly to health and well-being. In ethnic minority families, informal support has served to counteract the overall deleterious effects of adverse environmental, social, and economic conditions.

5.2. Sources and Types of Informal Support/Care

Studies that examine the beneficial effects that care networks have on health and well-being find that support derives from many sources, including spouse, children, siblings, and friends. Racial and ethnic minority families, in comparison with White families, tend to have larger multigenerational households. On the one hand, investigators report that generations live together out of economic necessity due to low socioeconomic status, poverty, and generally few resources and opportunities available to ethnic minorities. On the other hand, some researchers assert that intergenerational households may be more the product of cultural norms such as familism, filial piety, and familial obligation. In any case, the large family provides a functional, mutually beneficial network from which to give and receive care. Common characteristics of ethnic minority families are the importance of strong family bonds, exchanging of resources, and caring for the elderly.

As they age, many elderly individuals perceive that family and friends can be relied on to help with increasing needs. Family, friends, neighbors, and fictive kin play an important role in the care of ethnic minority elders. Mutual assistance, exchanges, and reciprocity involving informal and formal exchanges characterize the support patterns. A variety of tangible and emotional resources are exchanged within racial and ethnic minority populations, including goods and services as well as financial, emotional, and affective support. Instrumental aid (e.g., food, money, transportation) and emotional support (e.g., advice, counseling, visiting, companionship) are exchanged throughout the life course. These exchanges allow elderly individuals to remain viable and independent in their communities.

The elderly prefer assistance in a manner depending on need. Members of helping networks provide different types of support. Some members may give emotional support, such as listening and advice, whereas others may provide instrumental support, such as sick care and financial help. Family assistance may be more appropriate for long-term care, addressing impairments in activities of daily living (ADLs) and limitations in instrumental activities of daily living

(IADLs). Friends and neighbors, in close proximity to elders, may help with daily or short-term needs. Friends may also be a source of companionship and emotional support.

Research provides a wealth of information concerning formal and informal church support. The special role of church support is separate from its religious role for Black elderly individuals. Black churches have proven to be responsive to the needs of communities that have limited access to general societal support systems. Church members exchange instrumental, financial, emotional, and spiritual assistance with each other. This support includes food, clothing, and sick care as well as advice, encouragement, and information.

5.3. Reciprocity

Overall, older individuals are more satisfied with their relationships and have increased well-being if they are involved in reciprocal relationships. The nature of reciprocity exchanges is dependent on the nature of the relationships. Some relationships require immediate and in-kind or equivalent value return of services or goods. Other more intimate or lifelong relationships have more mutually satisfying exchanges over time and less immediacy to the exchanges. For example, receiving more aid than giving aid can result in feelings of guilt and indebtedness. One theory that addresses this issue is Antonucci's notion of a support bank or support reserves. The concept explains how many elderly individuals are able to maintain positive psychological well-being when they are in need of support but are unable to reciprocate. To illustrate, older minorities usually have fewer resources than do younger network members. By participating in years of giving and mutual exchanges, it becomes psychologically acceptable for the elderly to receive care without giving something in return. There is also evidence that some elderly individuals may reduce the size of their social exchange networks so as not to feel indebted to too many people. It is important to keep in mind that the elderly are not necessarily unable to render some type of reciprocity. For example, they might provide child-rearing assistance, including child care and advice, to younger parents in their networks.

5.4. Formal Care

At times, the elderly may need more professional care than family and friends can provide. When necessary, family, friends, and church members are instrumental in providing referrals to formal services such as professional home care providers. Ethnic minority elders also use formal services in the absence of spouses, children, or other informal helpers. Traditionally, the minority elderly have had few links to formal service. In the past, service delivery organizations rarely acknowledged cultural or family generational lifestyles differences, thereby perhaps contributing to mistrust of formal service organizations.

In general, the ethnic minority elderly are involved in long-term associations involving reciprocal support and care exchanges. More information is needed, however, about how societal changes affect these relationships. For example, younger people who become more educated and economically successful might need to move farther away from their family homes, thereby affecting the availability of care for elderly family members. Furthermore, people are living longer and having fewer children, and more middle-aged adult daughters are working outside the home. These changes also may contribute to a decrease in the number of people available for caring for elderly family members. Among recent immigrants, assimilation and acculturation issues also may complicate familial and friend relationships. The degree of assimilation varies and has effects on the provision of care. Despite histories of disadvantage, including poverty and poor medical care, ethnic minority groups have been creative at developing resources and effective coping strategies to alleviate these deleterious situations. Consequently, they have high levels of satisfaction with life, even in circumstances of relatively poorer health than that of Whites.

6. RECEIVING CARE

Although it has been noted that ethnic minority elderly have viable support systems, the helping literature suggests that receiving aid may have negative psychological consequences for care recipients who have health limitations. The logistical differences between receiving social support in general and receiving care for functional limitations are vague. Past research indicates that older Blacks have varied social support networks consisting of family as well as friends. But little research has examined the adequacy, satisfaction, or psychological implications of transitioning from self-sufficiency to dependency among older Blacks. The psychological effects of actual dependency differ from the psychological benefits of participating in

social support exchanges. Studies have focused on the psychological and physical effects of providing care to functionally limited elders, but little is known about care recipients' perceptions of the caregiving relationship.

It is often implicitly assumed that if care recipients with IADL limitations and ADL impairments have caregivers who adequately meet their physical needs, they are likely to be functioning satisfactorily both physically and mentally. In general, giving help is beneficial and should be encouraged; however, it is also complex and multifaceted.

Responses to aid depend on many factors, including the relationship and history of the donor with the recipient, the characteristics of the donor and the recipient (e.g., age, race, socioeconomic status, gender), and the appropriateness of aid to the needs of recipients (e.g., the context in which the aid is given). Research indicates that all of these factors can affect the recipient's sense of control. For example, if the help is threatening to feelings of personal control, it can lead to negative psychological outcomes for the recipient. However, if the help is supportive and appropriate to the needs of the recipient, it may reduce any threat to the recipient's concerns about control. Having some sense of control over who provides care and how much is provided may help to limit psychological distress.

Sense of control is a psychological resource that is effective in helping people to overcome environmental threats. Believing that one has control over negative events reduces possible adverse effects and promotes positive psychological outcomes. Control is indicative of psychological resilience and is effective in buffering the effects of stress. Perceptions of control have beneficial effects on individuals, whereas diminished control over undesirable events induces stress and anxiety and may lead to learned helplessness and depression. A sense of not being in control reduces motivation and adversely affects coping strategies, adaptation, and problem solving.

Some studies show an increase in control with age, whereas others demonstrate a decrease in control with age. The elderly usually encounter more loss and negative life events than do younger individuals. Events such as loss of one's spouse, physical limitations, and increased contact with health care providers are often associated with a decrease in control and adverse psychological outcomes. However, it is important to examine control within a specific context. Individuals do not feel the same extent of control over all sectors of their lives; degree of control varies from one aspect of

life to another. For example, control over career and social relations increases with age, whereas control over health and physical functioning decreases with age. How older people adapt to a new care-receiving situation may depend on caregiving factors noted earlier and control beliefs. Successful adjustment means better mental health outcomes. Believing that one has control over challenges and difficulties results in positive outcomes in various domains, such as psychological adjustment. Research has determined that the perception that an individual has control over his or her life is a stronger predictor of psychological well-being than is actual control. Thus, threats to perceived control in important domains of life may have particularly salient effects on overall health and well-being.

7. IMPORTANCE OF PERCEIVED CONTROL TO PHYSICAL AND MENTAL HEALTH OUTCOMES

Control has been studied in relation to a number of psychological constructs and has been found to be a strong predictor of physical and mental well-being. These constructs include life satisfaction, coping, depression, mortality, and morbidity. There is a strong relationship between helplessness (i.e., the opposite of control) and depression. Sense of control over both good and bad outcomes is associated with low levels of depression. Studies also indicate that people with a high sense of control are more satisfied with life and are more willing to face challenges than are those with a low sense of control. Attributing control to luck or chance fosters uncertainty and anxiety and may lead to depression. When people believe that they have a low or no degree of control, they become passive and withdrawn. People who believe that they have a higher level of control develop habits to solve and prevent problems.

Feelings of control are especially important for vulnerable populations. Sense of control not only affects psychological functioning but also affects the actions that people pursue as well as people's cognitions and emotions. For example, control influences how people characterize an event. Those who have low control expectations will attribute declining health to the aging process and, thus, do little to change or prevent further problems. Those with a higher sense of control will participate in intervention and prevention activities to improve their health. Elders receiving care due to physical impairments or health problems may be

feeling particularly susceptible to concerns about loss of control. Sense of control is an important construct in predicting adjustment among the elderly, particularly those with ADL impairments, such as problems with cooking, housework, shopping, and walking.

The infirm elderly may find themselves with care providers who prefer compliant care recipients who easily relinquish control in exchange for care. There may be expectations that care recipients should accept decisions made for them about who provides care and how much care is provided. These uncontrollable stressors, including physical limitations and problems in caregiving, may magnify the sense of loss of control and lead to depressive symptoms for elderly individuals receiving care. To illustrate, research suggests that when older people move into their children's homes, they often experience feelings of depression and increased helplessness and dependency, a diminished sense of personal control over their environment, and a loss of control over their destinies. However, when elderly people in care-receiving situations are given responsibility for some of their daily activities, they experience higher levels of control and fewer symptoms of depression and less dependency. Elderly people experience less psychological distress when they are able to control the course of events in their lives.

For people in poor health, a high sense of control is usually associated with high levels of psychological well-being. However, it should be noted that high perceptions of control have also been linked to poorer psychological well-being in people with health problems. This may be because people with a high sense of control and health problems may feel as though they are losing their independence. Specifically in the caregiving context, recipients with physical impairments may consider the assistance they receive to be intrusive. Nevertheless, the caregiving assistance is needed. Therefore, those who are impaired and have a high sense of control may be at greater risk for depressive symptoms than are those without a high sense of control. Traditionally, people with a low sense of control depend on others for help, whereas a high sense of control is related to greater motivation for making the decisions that affect people's own lives. Care receivers with a low sense of control may be better at adapting to relinquishing control of their daily lives to their caregivers. This relinquishing of control may result in positive psychological outcomes for care recipients as well as feelings of satisfaction with their relationships with caregivers. On the other hand, people with a high sense of control might feel more independent and perceive themselves as not giving up their autonomy and self-sufficiency in exchange for care. Consequently, these individuals might be at risk for poor care or, worse, for no care providers, eventually resulting in institutionalization.

Levels of sense of control may have unique effects on psychological well-being for ethnic minorities. Research suggests that because of past histories of discrimination, ethnic minorities might interpret some negative life events as uncontrollable. However, these interpretations are often associated with positive outcomes of psychological well-being. Racial and ethnic minorities may be well adapted to dealing with stressful negative life events and to acknowledging and accepting when they are unable to control or change the course of those events. When opportunities and resources are low, or when advantages are few, it may be healthier and more prudent to assume a low sense of control. Usually, an external sense of control is associated with unhealthy physical and mental health outcomes, but for the ethnic minority elderly, an external view may serve as a buffer to the effects of uncontrollable negative life events.

8. RECEIVING CARE, SENSE OF CONTROL, AND PSYCHOLOGICAL WELL-BEING

8.1. Caregivers

8.1.1. Burden, Stress, Depression as a Result of Providing Care

The caregiving literature reveals that spouses are the first choice for providing care. Spouses are willing and committed to helping each other, but often the outcome for caregiving spouses is depressive symptoms, anxiety, stress, and physical impairment. Spouses in caregiving situations may be at advanced ages with limited physical strength and stamina for the tasks of caregiving and sole responsibility of performing household tasks and home maintenance that were previously shared. The added responsibilities may lead to burden, stress, and depressive symptoms for caregiving spouses.

Adult children are also preferred over other nonrelative care providers. However, caregiving adult children usually have other obligations, including the needs of their spouses, children, and career goals. These other demands can be negatively affected by providing care to elderly parents. Some adult children providing care for parents often forfeit or delay career advancements due

to the responsibilities of caregiving. Furthermore, adult children may experience role reversal due to caring for their parents. Role exchange, career delay, and multirole responsibilities can have negative psychological consequences, such as role strain, depression, and poor physical health, for caregivers.

Studies comparing caregivers' psychological well-being indicate that Black caregivers report better psychological well-being and less negative psychological consequences of caregiving than do White caregivers. Investigators attribute these differences to cultural norms and beliefs and to patterns of living. For example, some studies show that Black caregivers report lower levels of caregiver burden and stress, and greater levels of caregiving mastery and satisfaction, than do White caregivers. Furthermore, ethnic minority caregivers often have less social and high-status career obligations to relinquish when assuming caregiver responsibilities than do Whites. As a result, they usually have less to lose economically than do their White counterparts.

8.2. Care Recipients

8.2.1. Effects of Caregiving on Care Recipients' Psychological Well-Being

Care recipients who must learn to master a new environment due to physical limitations or some other transition in their lives find that they would rather have the freedom to choose who helps, when they need help, or when they would rather do tasks themselves. When help is needed, recipients of aid are usually more willing to receive help from family and friends than from strangers. The more intimate the relationship with the helper, the better for the care recipient. Receiving help from family and friends permits care recipients to use little or no formal services and to remain in the community.

Although care recipients rely heavily on family, they also weigh the costs of time and effort to their caregivers. If the cost to the caregiver is low, the care recipient may be more comfortable in asking for or receiving help. If the caregiver is a spouse, it may be less costly than it would be for an adult child who has a family and other job and career responsibilities. However, some research has found that recipients' psychological well-being sometimes is negatively affected in caregiving contexts with both spouses and adult children. Other work has found that the elderly prefer to have less control, particularly during major

health-related difficulties. Giving up control to family or health care professionals can often be a relief if control is not possible. In most day-to-day living circumstances, the elderly desire to have a routine schedule, to be independent, and to make decisions for themselves.

The literature suggests that care recipients with spouses or other family care providers feel dependent on caregivers and depressed. In his research on receiving help, Newsom reported that recipients often have negative cognitive and emotional reactions to receiving care, including feelings of dependency, indebtedness, incompetence, and worthlessness. Furthermore, having family caregivers sometimes exacerbates those perceptions, especially if recipients perceive that family members are burdened, stressed, and depressed by providing care or if care is not appropriate for the IADL limitations (i.e., too much or too little).

8.2.2. Spouse Caregiver

The caregiving literature indicates that spousal care is less costly in terms of time constraints than is care from others. On the other hand, recipients with spouse caregivers report being more worried, depressed, and dependent than do recipients with younger relative caregivers. These negative effects can be attributed to concerns about overtaxing the elder caregiving spouse with additional responsibilities and tasks. The spouse receiving care may harbor a sense of guilt, realizing that he or she can no longer share the responsibilities of housekeeping tasks and other marital obligations and that the extra burden of caregiving and added household responsibilities could jeopardize the health of the caregiving spouse. In addition, research indicates that the quality of marital relations affects caregiving perceptions and psychological well-being for both spouses. Poorer quality relationships are exacerbated by caregiver burden and stress. These effects on the caregiving spouse may be manifested in the quality of care given and contribute to increased depressive symptoms for recipients with either a low or high sense of control.

Research shows that the quality of spousal care has a strong influence on care recipients' psychological well-being. However, the literature also presents evidence that even spousal care is not always helpful. Unhelpful care can lead to negative self-perceptions and to negative perceptions about caregivers. Spouse caregivers who over- or underestimate the abilities of care recipients can inadvertently criticize recipients'

recovery efforts. These criticisms erode self-esteem, decrease sense of control, and contribute to increased depressive symptoms in recipients. In general, spousal caregivers provide appropriate care; however, the effect of poor-quality care is more detrimental to mental health than high-quality care is beneficial to mental health. Inappropriate help can, at times, reduce the actual and perceived capabilities of recipients. Help implicitly threatens the sense of control of recipients, depending on the nature of the need and the help given.

8.2.3. Adult Children

Most care recipients have good relationships with their adult children caregivers; however, there are elderly individuals with adult children caregivers who perceive themselves as burdensome to their children. In a review of the caregiving literature within Hispanic communities, it was concluded that familial support systems for the impaired elderly could have detrimental consequences for both care recipients and adult child caregivers. The more dependent elders were on their children, the more depressive symptoms the elders reported. Depressive symptoms and dependency can be a result of role reversal and of the feeling of being a burden on children. Receiving support for shopping, cooking, and household tasks from children indicates that individuals are no longer self-sufficient, causing care recipients to feel as though they are losing control of their lives and their independence. Most elderly individuals would prefer to live independently from their adult children and to receive emotional support rather than instrumental support. Care recipients fear burdening their children and experience more psychological problems receiving help when adult children have to adjust their lives, including work schedules, to provide any type of support.

In addition, with family members performing chores and tasks for care recipients, time for social interactions with recipients may be reduced. After completing caregiving tasks, adult children may lack the time or energy to show affection or emotional support, resulting in elder recipients feeling lonely as well as burdensome to their children. Lack of affection can cause negative psychological consequences for recipients, particularly if the stress and burden that family members feel as a result of providing care is manifested in their behaviors, attitudes, and quality of care toward recipients. The elderly desire to continue familial relationships, but they also want to live independently.

8.2.4. Other Relatives, Friends, and Church Members

The literature indicates that the ethnic minority elderly in the community without family often have several sources of support. Studies have examined the similarities and difference in the structure, demographics, and living arrangements of elderly Black and White social and caregiving networks and found that elderly Blacks are more likely to have nonfamily and unpaid caregivers than are elderly Whites. Many elderly individuals may feel less of a burden to other relatives and friend caregivers than to spouses and adult children care providers. Nonfamily members may be providing care out of kindness and genuine concern, but the care might not be consistent, steady, or dependable. For example, neighbors or friends going to the market may offer to purchase groceries or prepare extra food, and church friends may provide emotional support and transportation. However, providing care on a daily basis, such as for bathing, dressing, and taking medication, might require more commitment than friends and neighbors are able to give. If help is required too often, care recipients may begin to feel burdensome to nonfamily members. Receiving help from nonfamily members may be distressing but necessary to overcome care needs. The familiarity found in families is usually not present with nonfamily caregivers. In addition, nonfamily members may begin to feel beset by too many requests for assistance.

Because of the need for help with physical limitations, such as eating, bathing, and walking, recipients with nonfamily caregivers might need more professional help than spouses or family members can possibly offer. Some daily tasks are more appropriate for trained formal caregivers. Having physical or medical needs met, however, does not guarantee that recipients are comfortable with the caregiving situations or with the caregivers. Elderly individuals who are reluctant to voice disapproval or who have no control (whether actual or perceived) over ineffective or excessive care are at risk for depression.

9. RACIAL AND ETHNIC MINORITY ELDERLY AS A RESOURCE

The information in this article is important for practitioners dealing with elderly care recipients. Frequently, psychological well-being, such as depressive symptoms in the elderly, are ignored or misdiagnosed, resulting in early morbidity and mortality among older persons

receiving care. Research indicates that some health care professionals believe that the symptoms of depression are inevitable in the elderly and focus more on physical and medical problems. This article suggests that elderly care recipients' psychological needs are also important and that certain aspects of the caregiving relationship are related, for example, to an increase in depressive symptoms. Unfortunately, most care recipients often feel that they should be appreciative of any help they receive and will not complain, even when the help is inappropriate and causes poor psychological well-being.

Consideration should be given to elder care recipients as a resource. The elderly often possess valuable knowledge, skills, and insights that they can share with others. The elderly generally have more positive perceptions about the quality of their lives when their relationships with family and friends are of a reciprocal or interdependent nature than when they are simply receiving help from others. Elderly individuals receiving care are not necessarily unable to reciprocate in the caregiving relationship. Elderly parents are more satisfied if they can exchange support with care providers. For example, care recipients often provide emotional and financial support to caregivers, serve as confidants and sources of advice, and provide child care and sick care. The relationships between caregivers and care recipients are found to be supportive and reciprocal for both parties when recipients are given some responsibilities. Recipients can help with child care for their adult children who work and provide caregiving. The elderly can also share their skills and knowledge with family caregivers and grandchildren. They can help teach younger children to read and write, and care recipients who have hobbies (e.g., knitting) can teach skills to younger generations, further helping adult children and strengthening family intergenerational bonds. The elderly often help to relate and preserve family heritage and serve as socialization agents. Moreover, although frail or disabled older adults may be dependent on family members for care and support, there is some indication that various caregivers are dependent on care recipients for economic and housing assistance. For example, when adult children are in the divorce process or are experiencing temporary unemployment, parents are usually their first source of help.

Caregivers should also be aware of the importance of the emotional needs of care recipients and how to provide necessary care in a nonintrusive way. Care recipients should have as much control as possible over who provides care. Caregivers should allow care recipients to become involved in the planning of their daily care. For example, recipients may help to create grocery lists to purchase foods of their liking and help to prepare meals of their choice. The literature indicates that when care recipients are allowed to make decisions about their routines and are given some responsibility for their well-being and the important events in their lives, their emotional and physical well-being improves. Recipients who must learn to master a new environment due to disabilities need the freedom to choose when they need help or which tasks they would rather perform themselves. Their perceptions of their needs may be different from their actual needs, but it should be possible to provide care in a way that is not perceived as intrusive, overwhelming, or useless. To provide the best possible experiences for both receivers and caregivers, practitioners should assess and evaluate emotional and psychological aspects, and not only physical and medical aspects, of caregiving situations.

9.1. Implications

One of the major implications of the literature in the area of caregiving and care receiving is that recipients can have poor psychological health even when caregivers are providing for their physical needs and IADL limitations. Recipients' psychological health is an important factor to consider when examining the caregiving environment.

As noted previously, research that has investigated care recipients and caregivers has largely addressed the medical and physical limitations of care recipients. It is erroneously assumed that if caregivers are available to help with those limitations, recipients are functioning satisfactorily. Contrary to expectations, being cared for by family for some limitations and by nonfamily members for others might not always be beneficial to recipients' psychological well-being. Sense of control may be related to decreased psychological well-being, such as depressive symptoms, and there is compelling evidence that receiving and giving care is multifaceted and complex, especially among racial and ethnic minority elders. The authors believe that more research is needed on recipients' psychological well-being, especially sense of control, within the caregiving relationship.

10. CONCLUSIONS

This article has reviewed some implications of research findings in the areas of socioeconomic status, social support systems, well-being, and family patterns,

focusing specifically on perceptions of control and family caregiving and care receiving. It discussed how these findings are important for an applied psychology of aging, particularly within a biopsychosocial model of human development and aging. As indicated at the beginning of the article, the population of the United States continues to grow, especially in racial and ethnic diversity. Along with this complex demographic growth are social, health, and well-being considerations that affect the structure and functioning of the entire society. A clearer understanding of the influence of ethnic, racial, and cultural dimensions on social and psychological processes of aging and human development is necessary to address this complicated change in the size, diversity, and needs of the expanding older population.

See Also the Following Articles

Cross-Cultural Psychology, Overview ■ Elder Caregiving ■ Health and Culture ■ Health Psychology, Cross-Cultural

Further Reading

Antonucci, T. C., & Jackson, J. (1997). The role of reciprocity in social support. In B. R. Sarason (Ed.), *Social support: An interactional view* (pp. 173–198). New York: John Wiley.

Jackson, J. S. (1993). Racial influences on adult development and aging. In R. Kastenbaum (Ed.), *The encyclopedia of adult development* (pp. 18–26). Phoenix, AZ: Oryx.

Jackson, J. S. (Ed.). (2000). *New directions: African Americans in a diversifying nation.* Washington, DC: National Policy Association.

Jackson, J. S., Chatters, L. M., & Taylor, R. J. (1993). *Aging in Black America.* Newbury Park, CA: Sage.

Lustbader, W. (1991). *Counting on kindness: The dilemmas of dependency.* New York: Maxwell Macmillan.

Rodin, J. (1987). Personal control through the life course. In R. P. Ables (Ed.), *Life-span perspectives and social psychology* (pp. 103–119). Hillsdale, NJ: Lawrence Erlbaum.

Taylor, R. J., Chatters, L. M., & Jackson, J. S. (1997). Changes over time in support network involvement among Black Americans. In R. J. Taylor, L. M. Chatters, & J. S. Jackson (Eds.), *Family life in Black America* (pp. 293–316). Thousand Oaks, CA: Sage.

Aggression and Culture

Michael Harris Bond

Chinese University of Hong Kong, Shatin, Hong Kong

GLOSSARY

aggression Acts of coercive control perceived as illegitimate for the social context in which they were enacted.

coercive control Influence applied to another person or group that harms the person's or group's physical, material, or social interests.

culture Ecological–social contexts that promote or restrain acts of interpersonal or intergroup control.

holocultural studies Research that treats a social unit, such as the nation, as the unit of analysis and that spans as much of the world as possible.

retaliation Application of coercive control against a person or group deemed to have attacked oneself or one's group unjustly.

This article conceptualizes aggression as coercive forms of social control, perceived as being applied illegitimately. Mundane social control is exercised through role prescriptions and is routinely met with compliance. Whenever social control is perceived to be exercised illegitimately, however, victims and their representatives respond to restore interpersonal and social order. That response may be retaliatory coercion. The type of response elicited will be shaped by personal factors such as the emotionality of the victim and the hedonic relevance of the anti-normative action; target factors such as the gender, status, and ethnicity of the perpetrator; social factors such as the degree of interpersonal support for retaliation or harmonizing; historical factors such as the strength of antagonistic ideologies among ethnic groups; economic factors such as the degree of relative inequality between the actors and between their social groups; and cultural factors such as the presence of honor or restraint codes. For the purpose of understanding aggression, culture may be understood as that complex set of external conditions that sustains or retards the development of those social psychological conditions that predispose toward interpersonal violence. This article elaborates these societal conditions and their psychological imprints on social actors.

1. CULTURE AND AGGRESSION: CONTEXTS FOR EXERCISING COERCIVE CONTROL

Cultures may be construed as ecological–social contexts that promote or restrain acts of interpersonal or intergroup control. These acts function to influence other persons or the groups they constitute to provide resources desired by the influencing agent while minimizing the costs required for their acquisition. The party being influenced is also acting to maximize benefits and

minimize costs while engaged in this dance of interdependencies. So long as this exchange of influence proceeds within normatively accepted guidelines, the process is regarded as an acceptable negotiation. Once one of the parties violates the other's norms of acceptable procedure or outcome, the other's influence attempts may shift and become more coercive.

Coercion in influence is characterized by the use of punishments or by the withholding of rewards from the target as a consequence of his or her behavior. It is these tactics of contingent control that are closely monitored in society because their use may violate social norms for exercising social influence over another's interests. Any escalation in the strength of influence tactics will typically be construed as intended by the other party but unethical, illegitimate, unjustified, savage, immoral, or wrong and will typically be met with a counterattack. One party triumphs, at least in the short term, leaving the other party subjugated and resentful, immobilized and hopeless, or dead and past harboring plans for revenge.

Harmony or at least a working truce may be restored between these interdependent parties. If that harmony has not been restored in ways that satisfy the aggrieved parties, however, memories of unrequited abuse will fester, motivating acts of revenge for past wrongdoings. These acts of retaliation may erupt if the social calculus shifts through changes in the balance of power, economic recession, environmental degradation, and the like.

In the course of this familiar drama, the label "aggressive" will often be deployed to describe antinormative tactics of influence exercised by the other party. In general, such acts result in physical or material harm to the other party, such as wounding or removal of possessions, but can also include symbolic impositions, such as insults or threats. However, these same acts are generally not construed as aggressive by those who perform them. Such dramatically different accounts of hurtful acts given by perpetrators and by victims have been well documented. Instead, if called to account by an accepted authority, "aggressors" will justify their use of coercive control in terms of self-defense, doing their job, retaliating for past injustices—that is, getting even; serving their group, organization, or nation; protecting the social order; and the like. Aggression, then, is not a characteristic of the action itself but rather a labeling process arising out of the social context within which it occurs. That label constitutes a disapproving judgment passed on the process by which social influence over another person is exercised. Homicide, robbery, rape, and assault are universally regarded as acts of aggression. Defending oneself against attack, performing painful surgery, executing a convicted rapist, and spanking a disobedient child are generally not so regarded because those acts are often socially approved.

Of course, the exercise of illegitimate tactics of control does not always result in retaliation by the aggrieved party. The target may instead undertake proxy control, influencing third parties—be they friends, family, or social authorities—to restrain, distract, or attack the "aggressor." Or recipients of such resented manipulation may withdraw from the relationship if an exit option is possible and if alternative relationships providing the same desired goals are available. Alternatively, they may accept such "altercasting" into a subordinate status, thereby acceding to a relationship in which the dominant party may exercise such tactics of control.

The use of any such response to counternormative control depends on the skills and resources available to the target. The target must be aware of alternative responses because not everyone has the social training and intelligence to appreciate alternatives other than retaliation. The target must be willing to use the alternative responses because some may be proscribed by socialization or normatively sanctioned by proximal social groups. The target must be able to mobilize the personal resources and the social agents necessary to exercise that form of influence because not everyone commands the required talents, time, energy, resources, and social credits with others to motivate their intervention.

2. CULTURE AS CONTEXTS FOR INFLUENCE

Culture is a polysemous construct of kaleidoscopic richness, variously defined by practitioners of different social sciences. It may be defined for purposes of doing psychology as a set of ecological–social constraints and affordances potentiating some behaviors, retarding others, or even rendering them irrelevant. This general psychological definition culture was elaborated by Bond as

> a shared system of beliefs (what is true), values (what is important), expectations, especially about scripted behavioral sequences, and behavior meanings (what is implied by engaging in a given action) developed by a group over time to provide the requirements of living (food and water, protection against the elements, security, belonging, social appreciation, and the exercise of one's skills) in a particular geographical niche. This shared system

enhances communication of meaning and coordination of actions among a culture's members by reducing uncertainty and anxiety through making its members' behavior predictable, understandable, and valuable.

This definition focuses on the socialized psychological "software" constituting the operating system for functional membership in a group setting. That group may be any social unit—a family, a work group, an organization, a profession, a caste, a social class, a political state, and the like. The preceding definition is silent about the characteristics of those social units that predispose toward the development of different types and levels of the shared software called culture. Nor does this definition focus on the types and levels of behavior likely to characterize that social unit and its members. It is precisely those contexts and consequences of culture that need to be elaborated whenever one addresses the link between culture and any behavioral outcome such as aggression.

The cultural factors relevant for explaining differences in aggressive behavior will depend on how aggression has been conceptualized. These factors would then be linked to mediating processes responsible for driving aggressive behaviors. Given the preceding discussion, relevant considerations for a cultural analysis would include institutional supports enforcing procedural norms for coordinating action and for dividing resources; norms about the legitimacy of claims made by outsiders or different others on one's resources, including ideologies of antagonism and social representations of a group's history, focused on specific out-groups and their members; direct or indirect socialization for coercive responses to particular targets (e.g., women, the disabled, other lower status persons); training provisions and social support for conflict resolution strategies other than retaliation; the goals that social group members hold about desirable material and social outcomes and the procedures for achieving those outcomes (i.e., their instrumental and terminal values); attitudes held by citizens about the proper place of out-groups (i.e., social dominance orientation); and beliefs about the efficacy of various influence tactics, citizen neuroticism, and emotionality affecting impulse regulation. Each of these mediators is made salient in the issue of culture's relation to aggression by its hypothesized role in eliciting coercive tactics of interpersonal and social control. Whether a person or group using such coercive violent means of social control would be perceived as behaving aggressively would depend on the position of the actor and the target in the social order of their culture.

3. HOLOCULTURAL STUDIES OF AGGRESSION

Dramatic and universally accepted acts of aggression, such as homicide and political violence, can be studied only at the level of larger social units, such as societies and nations, because their incidence is either rare or group encompassing. Such studies are called "holocultural" because they involve social units sampled from the whole range of the earth's cultures. Aggression-relevant data on larger social units are easier to obtain than are individual-level data on cross-cultural comparisons of coercive control. Perhaps this is one reason why there is so little comparative psychological work on aggression despite its social importance. The problem with holocultural studies, however, is that it is more difficult to extract defensible social psychological explanations for their results.

3.1. Waging War and Internal Political Violence

In 1994, Ember and Ember examined the ethnographic atlas for clues about the causes of war in 186 societies. They concluded that societies wage war due to recurrent natural disasters that rendered them vulnerable to food shortages: "Fear of such unpredictable shortages will motivate people to go to war to take resources from others in order to protect against resource uncertainty." The societies they studied were smaller than current nation-states, which are enmeshed in relations of trade interdependence and protected by mutual defense pacts. These state-level connections act to restrain war, although the driving force of resource acquisition remains a powerful incentive for waging war. Looming environmental disasters, with their potential for reducing current national resources such as arable land and potable water, argue for carefully coordinated international programs of disaster relief and foreign aid. In their absence, pressures for waging war will mount as some nations suffer reversals of fortune.

In 1988, Rummel assessed war making during the course of the 20th century. He concluded that the historical record shows that democratic political systems—that is, "those that maximize and guarantee individual freedom"—are less likely to engage in war. Rummel argued that "there is a consistent and significant, but low, negative correlation between democracies and collective violence" and further that "when two

nations are stable democracies, no wars occur between them." Rummel argued that this relationship is causal and not merely correlational, stating that a democracy "promotes a social field, cross-pressures, and political responsibility; it promotes pluralism, diversity, and groups that have a stake in peace." Institutional pressures and supports in a democracy, from the opportunity to vote anonymously to legal supports for human rights, lead to the development of psychological processes, such as egalitarian values and trust in others, that prevent political leaders from mobilizing the mass support required for waging war. The same argument applies to restraining internal political violence. In this regard, Rummel provided this sobering reminder:

> War is not the most deadly form of violence. Indeed, while 36 million people have been killed in battle in all foreign and domestic wars in our century, at least 119 million more have been killed by government genocide, massacres, and other mass killing. And about 115 million of these were killed by totalitarian governments (as many as 95 million by communist ones).

He hastened to point out that during the 20th century "there is no case of democracies killing en masse their own citizens."

The preceding discussion suggests that if the big picture on aggression is considered, concerted efforts toward educating populaces for effective democratic citizenship would be the most powerful prophylactic against violence and investment in social capital that the leaders of a nation could make. Of course, the courage and insight required for nondemocratically installed leaders to commit the resources necessary for this task and imperil their own power basis will be hard to find without external political influence being engaged.

3.2. Homicide

The most reliable data on levels of homicide in a social unit are provided by the standardized measures collected by the World Health Organization, although judgments of homicide levels in a society have also been taken from the anthropological record provided by the Human Relations Area Files. These homicide levels are typically correlated with societal factors to provide insights about characteristics of social systems that predispose members of their citizenry toward lethal aggression against one another. Such findings are springboards for some social psychological explanation of violence by citizens against fellow citizens. Occasionally, empirical data are

provided to support such speculations about social psychological processes mediating these acts of violence.

3.2.1. Societal Factors

Most research in this area is bivariate, where one societal predictor, such as national wealth, is correlated with the rate of national homicide. Over time, a number of societal factors have been advanced as "causing" higher levels of internal murder. Theory is then advanced to explain each variable's avenue of influence. One problem with this piecemeal approach is that many of these factors are interrelated and their avenues of influence are overlapping or perhaps identical. Understanding the full range of societal relations to homicide rate requires multivariate studies where a host of potential societal predictors is deployed, so that their separate influences may be assessed. Such studies could then confirm many of the past findings, integrating theorizing around key processes such as social trust. It could also add new predictors to the societal equation, suggesting further lines of theoretical development.

It is possible to use these fewer multivariate studies to integrate the results from the various studies that have been conducted on homicide rates over the past 40 years. Using this approach, the following societal variables have been linked to homicide rate: a society's recent history of waging war, its level of wealth (negatively), its degree of economic inequality, its percentage of male unemployment, its strength of human rights observance (negative), and a measure of male dominance themes in its national system (labeled "female purity"). More studies have been conducted using the readily available economic indexes, so one can be more confident in their robustness of impact, especially that of relative inequality. The other variables are new entrants in the predictive arena, emerging as more national indicators become available for use by social scientists during the 21st century.

3.2.2. Social Psychological Mechanisms

Investigators of societal variables related to homicide have often used their empirical results to support the probable operation of certain social psychological processes in potentiating homicide. So, for example, Wilkinson argued in 1996 that relative economic inequality predisposes the "have-nots" in a social system to feel routinely status deprived and inferior. The social stress so induced increases the likelihood of violent behavior in everyday social exchanges. Taking a complementary approach, Fukuyama argued in 1995 that the

higher levels of incivility, including homicide, in poorer societies arise through lower levels of social trust held by citizens in such polities. This higher ambient trust among a society's members may be sustained by greater equality in resource distribution because the two features of a social system are highly correlated.

In their 1994 analysis, Ember and Ember did more than speculate on the social psychological processes involved in homicide; they empirically demonstrated that the societal link between a recent history of warfare and a society's internal homicide rate is mediated by the stronger socialization of males for aggression in such societies. This demonstration is accomplished by using regression techniques that enable the interrelatedness of societal and social psychological variables as predictors of homicide to be assessed within the same data set. The richness of the linkages so revealed, however, depends on the types of social psychological variables included in the data set and their overlap with the nations whose homicide rates are being studied. Because these data sets consist of nation scores, considerable data collection must be undertaken to provide the necessary number of "citizen" averages on the social psychological predictors. This multicultural database is becoming more and more available with the increasing activity in cross-cultural psychology, so the linking of societal factors to psychological dispositions of a citizenry to national rates of violent outcomes is becoming possible.

Using one such database, Lim and colleagues found that national rates of homicide across 56 nations were jointly predicted by a nation's level of economic inequality, its gross national product (GNP) growth over a decade (negative), and its emphasis on female purity (Fig. 1). This finding confirmed the results of many previous studies of economic inequality and homicide but extended them to include additional societal predictors of internal violence. Lim and colleagues also found that citizen scores on the psychological constructs of belief in fate control and the length of typical emotional experiences predicted national rates of homicide. Using regression techniques, they concluded that the effect of female purity was mediated by citizen endorsement of fate control, that is, the belief that one's outcomes were shaped by one's predetermined destiny. This linkage led to speculation that in societies characterized by social divisions and rigidity of social structure, parents typically socialize their children for responsiveness to external control rather than for internal self-regulation. Lacking strategies for internally generated self-control, citizens in such societies are more likely to respond with homicide to the universal frustrations arising out of mundane interpersonal interdependencies.
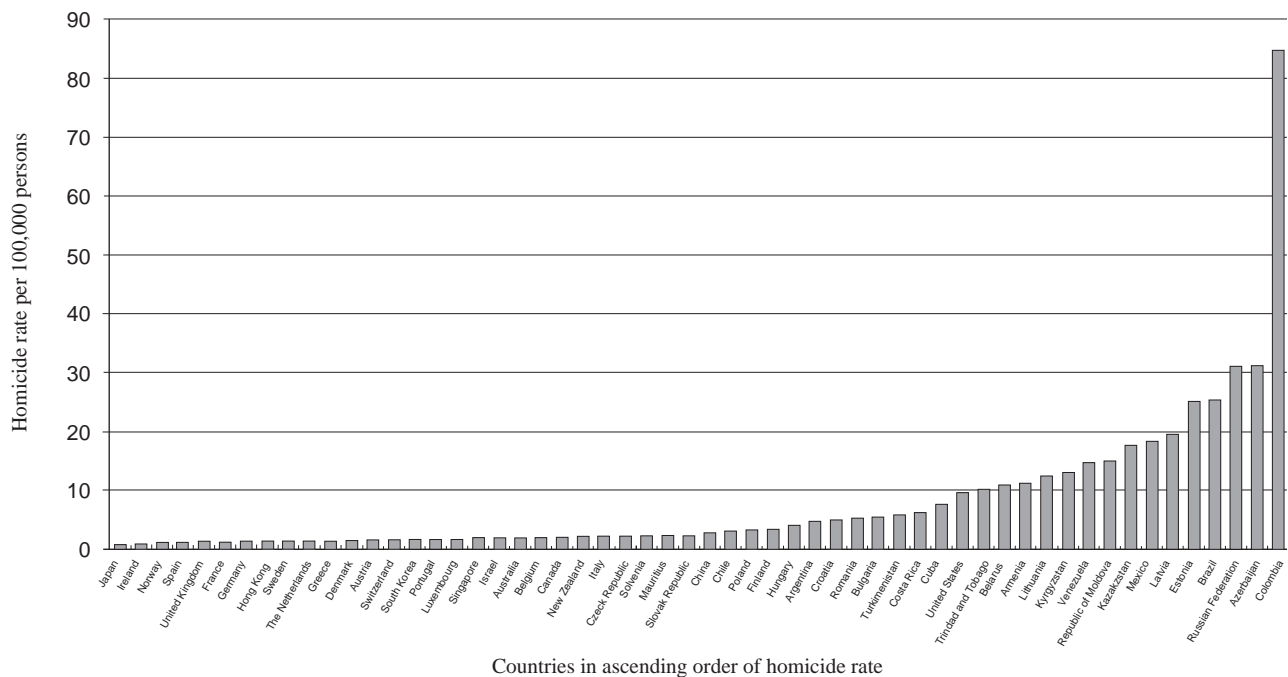

FIGURE 1   Homicide rates in 56 countries.

In nations where citizens routinely experience emotions for longer periods of time, there would be a persistence of emotionality into more of one's daily interactions with others. This wider reach of emotionality would amplify whatever dominant tendencies characterized a particular exchange. It is easy to understand how, if that exchange involved coercive control or resistance to its imposition, the frequency of homicide as one extreme outcome would be higher in such countries.

Such conclusions are speculative because they are drawn from correlations of national-level scores. In the case of homicide, however, there is little choice but to use data from large social units. The plausibility of the conclusions drawn will depend on their connectability to theory and empirical data derived from analyses of individual behaviors related to homicide. For example, a person's capacity for self-regulation has been linked longitudinally to lower frequencies of antisocial acting out. So, it is plausible that belief in control by fate is part of a personality constellation related to poor impulse control. Psychological studies of individuals differing in their beliefs about fate control will be necessary, however, to strengthen the plausibility that citizen differences in fate control link to national homicide rates through the agency of impulse regulation.

4. CULTURE AND STUDIES OF AGGRESSION IN INDIVIDUALS

The holocultural studies discussed previously alert social scientists to the need for validated theories about the psychological mechanisms involved in generating violent behavior against others. These mechanisms must be measurable across cultural systems, and their role in potentiating harm to others must be demonstrated empirically. This is the mandate for cross-cultural psychology, but its offerings are remarkably scanty. In part, this shortfall arises due to the definitional problems noted earlier (e.g., what behaviors other than homicide, rape, or physical assault constitute aggression?) There are a plethora of American laboratory studies, for example, that use paradigms such as the teacher–learner shock scenario to examine aggression. In such a setup, the participant-teacher trains her confederate-student under various experimental conditions hypothesized to provoke higher levels of hostility (e.g., ambient temperature, expectation of future interaction, prior insult by the confederate-student). This procedure, however, fails to engage the dynamics of aggression as anti-normative

behavior because the experimenter has explicitly legitimized the delivery of noxious stimuli to the student as an educational exercise. Other experimental scenarios for studying aggression and its facilitating conditions likewise encounter the same problem of its implicit legitimation by the experimenter.

A related problem for the cross-cultural laboratory study of aggression involves the meaning of the exchange between the two parties created by the experimenter. In the teacher–learner scenario, one could imagine that in more hierarchical cultures with a tradition of corporal punishment in classrooms, the legitimation of hurtful behavior by the experimenter would appear to be more acceptable to the participant-teacher. The confederate-student would then receive even higher levels of shock for a given set of experimental conditions. One could not, however, conclude that persons of this hierarchical culture were more aggressive given that local understanding of the shocking behavior is so different in the new culture.

Given these conceptual problems, two ways forward have been tried. One approach across cultural groups involves the use of scenario studies where cultural informants are presented with descriptions of behavioral exchanges between persons and are asked to make judgments about these exchanges and the persons involved. Knowledge about the cultures can then be used to predict the outcomes of these judgments. So, for example, Bond argued that in a more hierarchical culture, a direct coercive tactic of influence in a business meeting would be construed differently than it would in an egalitarian culture, depending on its source. As predicted, a superior insulting a subordinate was less sanctioned in Chinese culture than in American culture, and a subordinate in Chinese culture was more strongly sanctioned for insulting a superior than was a subordinate in American culture for insulting a superior. Similarly, knowledge about cultural dynamics can be used to predict the frequency and judged effectiveness of direct coercive tactics of influence, depending on their source and their relationship with the target. In such studies, the behavior being judged is the same, and its different meaning is examined in light of different cultural dynamics. The perceived aggressiveness of the behavior and the sanctions it receives vary as a function of these dynamics.

The second approach involves observation of hostile behavior in common settings such as prisons and schools. These observational studies are time-consuming and require agreement on what behaviors to measure. There has been considerable debate across participants

in the European Union, for example, as to what behaviors constitute "bullying," so that cross-national comparisons can proceed. Typically, a "least common denominator" position that includes only physical or verbal provocation between perpetrator and victim is adopted. Other indirect forms of bullying, such as gossiping and cliquing against others, are excluded. This happens in part because persons from different cultural traditions might not regard these behaviors as bullying, but also because the indirect behaviors are extremely difficult to measure reliably. Given all of these concerns, few cross-cultural studies involving *in situ* measures of behavior have been conducted to date.

Less conceptual and measurement difficulty is encountered when behaviors are measured by asking knowledgeable observers, such as parents and teachers, to rate the frequency with which their charges engage in aggressive behaviors. Chang, for example, employed a 7-item scale measure of aggressive behaviors such as hitting, teasing, and shoving others. The scale forms a coherent single measure of aggressiveness, and independent observers of a given pupil rate that pupil's aggressiveness with high degrees of agreement. It would be easy to examine this measure in other cultural contexts for its equivalence, adding or subtracting behaviors until a comparable measure of pupil aggressiveness is achieved. Cultural groups could then be compared for their levels of aggressiveness.

There is a slowly growing body of cross-cultural research using the scenario and behavior-reporting approaches. The problem with this work is that it is often only descriptive, reporting the cultural differences and perhaps speculating why they occur. There is little theory-driven work that measures the processes hypothesized to drive the aggressive behavior in the cultural groups studied. One exception is a recent study by Chow and colleagues using recollections by cultural informants of upsetting interpersonal events. They argued that everyday aggressive behavior often emerges out of social exchanges where one party harms another party, resulting in retaliation from the aggrieved party. Self-reports of behaviors undertaken by the victim after the harm doing indicated an equivalent cluster of assertive or aggressive behaviors in both Japanese and American informants. This assertive counterattack is driven by psychological factors uniting to generate a motive to retaliate and by social factors relevant to the relationship such as the gender of the two parties and their degree of familiarity.

Chow and colleagues found that both internal psychological and external social factors contributed toward assertive counterattack in both Japanese and American cultures but that psychological factors weighed more heavily than social factors for the Americans than for the Japanese. This finding confirmed theoretical speculation that collectivists, such as Japanese, are more responsive to external forces, whereas individualists, such as Americans, are more responsive to internal forces. This reasoning was supported by personality measures of individualism that predicted the tendency to emphasize psychological factors more than social factors in directing one's retaliation against the harm doer. Controlling for the level of harm experienced, the investigators found that Americans were more likely to counterattack assertively because their motivation to retaliate was greater than was the Japanese motivation.

The next step in such research on culture and aggression is to examine why the motivation of any cultural group to counter any harm with coercive retaliation is greater than that of another cultural group. One strong possibility is the experience of harsh parenting. Punitive parents model coercive strategies of control as they manage their children, indirectly teaching children to respond to interpersonal frustration with confrontational assertiveness. This linkage between harsh parenting and children's aggressiveness has been confirmed in other settings such as school. Such studies have provided the inspiration for successful intervention studies that reduce the level of child aggression through programs teaching parents to intervene nonviolently.

Chang further discovered that some of the linkage between harsh parenting and child aggression is mediated by lower levels of affect regulation characterizing those pupils who act out. Such internal dispositions may be higher in citizens of some cultural groups than in citizens of other cultural groups due to differential socialization styles and procedures. The different strengths of these aggression-related attributes could then account for observed cultural differences in aggressive responding. Other possible internal mediators include well-studied psychological dispositions such as self-regulatory efficacy, sociopathy, and state anger. Social psychological factors varying across cultures could include orientation toward confrontation, honor sensitivity among males, degree of ethnic identification, social support for retributive as opposed to restorative procedures of justice, and socialization into ideologies of antagonism and indebtedness.

Higher cultural levels of these psychological and relational dispositions may be the legacy of cultural circumstances (e.g., a history of war, tribal conflict, or colonialism), a herding economy, slavery, or internal political violence, any of which may lead parents to

socialize their children to inflict harm on others as a way of solving problems in interpersonal and intergroup coordination. As indicated earlier in the discussion of holocultural studies, psychologists have some ideas about what these cultural circumstances might be. But they need to conduct the sort of theory-driven and empirically validated research necessary to substantiate their speculations about how culture operates to affect aggression. Psychologists must instrument the hypothesized dispositions, measure them in persons, and then observe their behavior when dealing with interpersonal and intergroup relationships. Who chooses cooperation and consensus? Who chooses confrontation and conflict? How then can life be enhanced by encouraging social systems to socialize their citizens for the former rather than for the latter?

See Also the Following Articles

International Conflict ■ Interpersonal Conflict

Further Reading

Bond, M. H. (2004). Culture and aggression: From context to coercion. *Personality and Social Psychology Review*, *8*, 62–78.

Chang, L. (2002). *Emotion regulation as a mediator of the effect of harsh parenting on child aggression: Evidence from mothers and fathers and sons and daughters*. Unpublished manuscript, Chinese University of Hong Kong.

Chow, H., Bond, M. H., Quigley, B., Ohbuchi, K., & Tedeschi, J. T. (2002). *Understanding responses to being hurt by another in Japan and the United States*. Unpublished manuscript, Chinese University of Hong Kong.

Dishion, T. J., Patterson, G. R., & Kavanagh, K. (1992). An experimental study of the coercion model: Linking theory, intervention, and measurement. In J. McCord, & R. Trembley (Eds.), *The interaction of theory and practice: Experimental studies of interventions* (pp. 253–282). New York: Guilford.

Ember, C. R., & Ember, M. (1994). War, socialization, and interpersonal violence: A cross-cultural study. *Journal of Conflict Resolution*, *38*, 620–646.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York: Simon & Schuster.

Lim, F., Bond, M. H., & Bond, M. K. (2002). *Social and psychological predictors of homicide rates across nations: Linking societal and dispositional factors*. Unpublished manuscript. Chinese University of Hong Kong.

Rummel, R. J. (1988, June). *Political systems, violence, and war*. Paper presented at the U.S. Institute of Peace Conference, Airlie, VA. (Available: www.shadeslanding.com)

Segall, M. H., Ember, C. R., & Ember, M. (1997). Aggression, crime, and warfare. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology* (Vol. 3, pp. 213–254). Needham Heights, MA: Allyn & Bacon.

Sidanius, J. (1993). The psychology of group conflict and the dynamics of oppression: A social dominance perspective. In W. McGuire, & S. Iyengar (Eds.), *Current approaches to political psychology* (pp. 183–219). Durham, NC: Duke University Press.

Staub, E. (1989). *The roots of evil*. Cambridge, UK: Cambridge University Press.

Tedeschi, J. T., & Bond, M. H. (2001). Aversive behavior and aggression in cross-cultural perspective. In R. Kowalski (Ed.), *Behaving badly: Aversive behaviors in interpersonal relationships* (pp. 257–293). Washington, DC: APA Books.

Tedeschi, J. T., & Felson, R. B. (1994). *Violence, aggression, and coercive actions*. Washington, DC: American Psychological Association.

Wilkinson, R. G. (1996). *Unhealthy societies: The afflictions of inequality*. London: Routledge.

Aging and Competency

Jennifer Moye

Brockton VA Medical Center, Boston, Massachusetts, USA

GLOSSARY

autonomy A concept in biomedical ethics that refers to an individual's right to self-determination and to make decisions and choices concerning one's own life.

beneficence An ethical principle important to consider in guardianship issues that refers to the obligation to intervene to positively benefit another individual and prevent harm.

capacity Another word for competency referring to the ability to receive and evaluate information and to make and communicate decisions.

competency A legal status presumed for all adults—as opposed to incompetency, wherein a person with a clinically diagnosed condition is unable to receive and evaluate information or to make and communicate decisions to such an extent that he or she lacks the ability to meet essential requirements for health, safety, and self-care; judicial determinations of incompetency are not necessarily permanent and may be revoked if and when the ward's mental status changes.

conservatorship A legal mechanism established by a court after a hearing that empowers one party to make financial decisions (e.g., management of assets, businesses, contracts, wills, and gifts) for another party.

durable power of attorney (DPA) An alternative to guardianship in which a competent individual designates another person to manage his or her affairs (medical and/or financial) in the event of incapacity.

guardianship A legal mechanism established by a court that designates a party to make decisions (e.g., health, financial, residence) for another (incompetent) individual.

health care proxy Any person appointed by another competent individual to have the authority to make health care decisions if and when he or she becomes incapable of making such decisions; the proxy's authority may be unspecified (to make all decisions in accordance with the principal's best interests) or may be specified (to make decisions in accordance with advance directives).

representative payee An individual appointed by a benefit provider (e.g., veterans disability, social security disability) when the provider questions the benefit recipient's ability to manage the funds; the payee receives and has responsibility for only the funds distributed by that provider.

Many people use the concept of competency in clinical settings, although there continues to be confusion about what it means, when it is relevant, and how it should be assessed. All adults are presumed to be legally competent unless otherwise determined in a court of law. An adult's competency may be questioned when a family member or other interested party files a petition for guardianship of that individual. If incompetency is found in a court hearing, a guardian or conservator may be appointed. In most states, a guardian is responsible for decisions regarding the care of the individual's "person" (e.g.,

day-to-day issues of care, residence, and activities), whereas a conservator is responsible for decisions regarding the care of the "estate" (i.e., management of assets).

1. CLINICAL VERSUS LEGAL COMPETENCY

Although clinicians may question a patient's competency status and may request other clinicians to evaluate the patient's competency, a clinical finding of incompetency should not be confused with a judicial determination of incompetency. The clinical use of the term "competency" refers to a clinical opinion regarding the patient's decisional capacities. Such an opinion may be used to activate a durable power of attorney or health care proxy if one exists. However, clinical statements about capacities should not be equated with court-determined findings of incompetency.

2. COMPETENCY VERSUS CAPACITY

Increasingly, the term "capacity" is preferred to competency. The use of capacity recognizes the possibility of individual strengths and weaknesses or functional capacities rather than the all-or-none notion of incompetency.

3. LEGAL FRAMEWORKS FOR AGING AND COMPETENCY

In the United States, most statutory definitions of incapacity to care for one's person or estate include four parts, as shown in Fig. 1. This four-part definition means that a diagnosis is a necessary, but not a sufficient, component of incompetency. Rather, information must be provided as to how the disease affects attention, memory, information processing, and so forth as well as, in turn, how these symptoms affect the person's ability to do certain things or make specific decisions. The ability level should be considered within the context of the environmental demands and resources as well as specific situational risks and benefits. For example, the threshold for the capacity to reasonably manage a small amount of day-to-day cash spending would be different from the capacity to manage a complex portfolio of investments, and the abilities needed for the capacity to consent to a relatively harmless medical intervention (e.g., a flu shot) would be much lower than those needed to consent to a more risky medical intervention (e.g., heart bypass surgery).
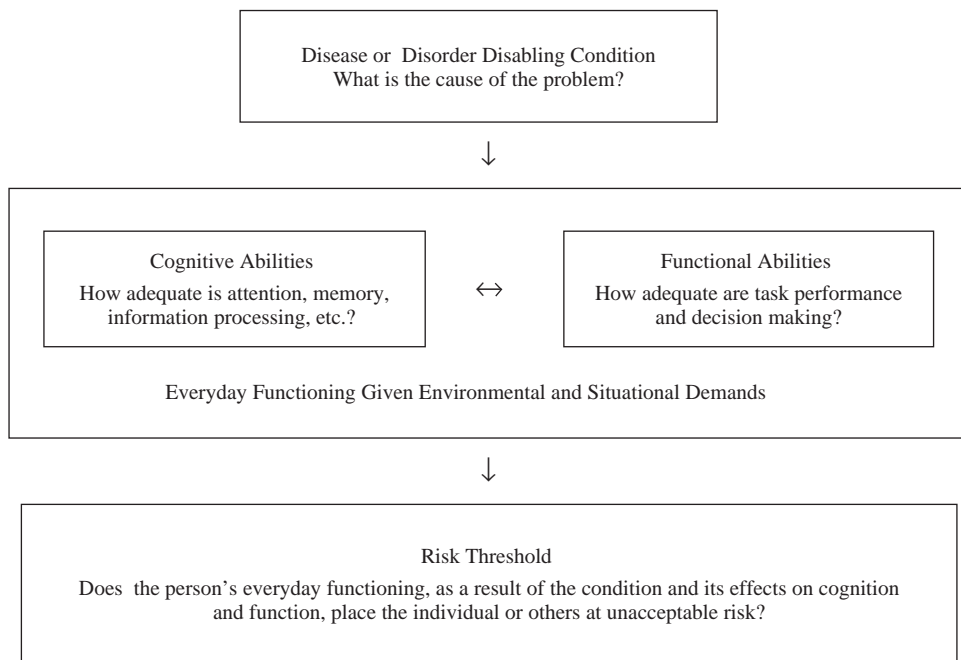
FIGURE 1   Four components of most U.S. statutory definitions of competency.

4. DECISION-MAKING ABILITIES IN COMPETENCY

Much U.S. case law about competency to consent to treatment refers to four decision-making abilities, as shown in Table I. Case law typically refers to these abilities individually as independent legal standards. Together, they form a useful framework for evaluating decisional processes.

5. CLINICAL EVALUATIONS OF CAPACITIES

5.1. Competency and Older Adults

Although the large majority of guardianships concern elderly individuals with psychiatric or neurological diagnoses, it is clear that advancing age or physical frailty themselves are not grounds for guardianship. Competency questions are increasing for older adults due to the age-associated increase in prevalence of dementia that can affect competency during the later stages of the disease.

TABLE I
Four Decision-Making Abilities Important for Legally Competency

Capacity	Definition
Understanding	The ability to comprehend diagnostic and treatment-related information and to demonstrate that comprehension
Appreciation	The ability to determine the significance of treatment information relative to one's own situation, focusing on beliefs about the actual presence of the diagnosis and the possibility that treatment would be beneficial
Reasoning	The process of comparing alternatives in light of consequences through integrating, analyzing, and manipulating information; also defined as providing "rational reasons"
Expressing a choice	The ability to communicate a decision about treatment that applies to individuals who cannot express a choice and also to those who are consistently ambivalent

5.2. Competency Assessment

In the past, most clinical evaluations of competency relied on a physician's interview or knowledge of the patient. However, such approaches can be unreliable. A more objective evaluation of competency, particularly important in more complicated cases, involves three parts, as shown in Table II.

5.3. Ethical Issues

5.3.1. Consequences of Guardianship

Many adults who receive guardianship experience benefit from the arrangement. In the best case, the

TABLE II
Components of Competency Assessment

Component	Definition
Clinical interview	
Who	Patient and relevant informants
What	Patient's values, goals, and preferences
	Situational factors
	Mental health conditions
	Symptoms of other illness
How	Open-ended interview
	Psychiatric rating scales
Structured assessment of cognition	
Who	Patient
What	Attention, memory, executive function, language, visuospatial abilities
How	Cognitive screening tests
	Neuropsychological tests
Performance based assessment of the specific capacity in question	
Who	Patient
What	Capacity to do the task or make the decision being questioned
How	Activities of daily living and instrumental activities of daily living instruments
	Capacity assessment tools

individual receives needed care, supervision, and advocacy in accordance with his or her wishes. In the worst case, guardians may take advantage of the individual's assets. In either case, removal of competency results in the loss of the right to make choices about residency, health care, medication, relationships, marriage, contracts, voting, driving, use of leisure time, and how one's own money is spent. Such losses in autonomy may affect the individual's emotional well-being. These potential rights deprivations and personal consequences of incompetency underscore the special care that should be taken in competency assessments and the extra attention that should be devoted to ethical concerns.

5.3.2. Informed Consent

An individual who is the subject of a competency evaluation needs to be informed of the nature of the evaluation and the potential consequences. Obtaining informed consent for the evaluation of competency from a person whose decision-making capacity is in question requires special care. Clinicians should fully disclose information about the evaluation and its risks and benefits in clear and direct language and then should assess the individual's ability to understand, appreciate, and reason through the information and options.

5.3.3. Individual and Cultural Differences

Persons from different ethnic or age backgrounds may have different cultural and cohort preferences for medical care, finances, living situations, relationships, and so forth. Evaluators must be sensitive so that individual and cultural differences are not misconstrued as incompetent decisional processes. This also applies to the use of standardized tests because some do not have adequate normative data concerning cultural, cohort, and educational influences on test performance, especially in older populations. Test instruments should be given in the primary language of the person being assessed.

6. CATEGORIES OF CAPACITIES

6.1. Financial Management

As shown in Table III, one important category of competency is the capacity to manage finances, including managing assets and how one spends money, managing debts and how one pays bills, and managing specific issues of contracts, disposition of property, and wills. Such financial capacities may involve knowledge of facts (e.g., where one's bank account is), financial skills (e.g., counting change), and financial judgment (e.g., avoiding

TABLE III
Categories of Competency

Category	Possible tasks	Potential services
Finances	Managing assets	Bill-paying services
	Paying bills	Money management services
	Writing will	
Health	Medical decision making	Visiting nurse
	Health care management (including medications)	Pill box and pill-dispensing systems
		Telephone reminder systems
Independent living	Household cleaning and maintenance	Homemaker services
	Laundry	Meals on wheels
	Meal shopping and preparation	Emergency call systems
	Communication	Home health aide
	Personal hygiene	Care management
		Adult protective services
		Assisted living
		Adult foster care
Transportation	Driving	Skill training classes
	Use of public transit	Rides to appointments and services
		Assisted public transportation

fraud). As indicated in the table, services can sometimes provide alternatives to guardianship.

6.2. Health Care

Another category of competency focuses on health care, especially the capacity to consent to treatment. Health care capacities may also cover managing day-to-day health such as nutrition, wound care, and medications.

6.3. Independent Living

A third and more broad category of competency concerns the capacity to live independently. Tasks in this category include household cleaning and maintenance, laundry, meal shopping and preparation, and communication (telephone and mail). Senior service agencies have a wealth of services to assist vulnerable adults in living independently, including homemaker services to clean, do laundry, and prepare meals; "chore" services to tackle larger tasks such as lawn care and snow removal; and meal delivery programs. Communication aids, including large-button telephones, assistive devices, and animals for the deaf and hard of hearing, may also promote independent living. Medical alert systems can be installed for seniors to enact in the event of a disabling emergency.

6.4. Transportation

A more narrow category of functional abilities concerns transportation and the capacity to drive a motor vehicle. This question may arise outside of other competency concerns when an adult with a disabling psychiatric or neurological condition has a series of motor vehicle accidents.

See Also the Following Articles

Cognitive Skills: Training, Maintenance, and Daily Usage ■ Decision Making ■ Elder Caregiving ■ Ethics and Social Responsibility

Further Reading

Department of Veterans Affairs. (1997). *Clinical assessment for competency determination: A practice guideline for psychologists.* Milwaukee, WI: National Center for Cost Containment.

Edelstein, B. (2000). Challenges in the assessment of decision-making capacity. *Journal of Aging Studies*, *14*, 423–437.

Earnst, K. S., Marson, D. C., & Harrell, L. E. (2000). Cognitive models of physicians' legal standard and person judgments of competency in patients with Alzheimer's disease. *Journal of the American Geriatric Society*, *48*, 919–927.

Grisso, T. (2002). *Evaluating Competencies* (2nd ed.). New York: Plenum.

Grisso, T., & Appelbaum, P. S. (1995). The MacArthur Treatment Competence Study III: Abilities of patients to consent to psychiatric and medical treatments. *Law and Human Behavior*, *19*, 149–174.

Marson, D. C., Ingram, K. K., Cody, H. A., & Harrell, L. E. (1995). Assessing the competency of patients with Alzheimer's disease under different legal standards. *Archives of Neurology*, *52*, 949–954.

Marson, D. C., McInturff, B., Hawkins, L., Bartolucci, A., & Harrell, L. E. (1997). Consistency of physician judgments of capacity to consent in mild Alzheimer's disease. *Journal of the American Geriatric Society*, *45*, 453–457.

Moye, J. (1999). Assessment of competency and decision making capacity. In P. Lichtenberg (Ed.), *Handbook of geriatric assessment* (pp. 488–528). New York: John Wiley.

Moye, J., & Karel, M. J. (1999). Evaluating decisional capacities in older adults: Results of two clinical studies. *Advances in Medical Psychotherapy*, *10*, 71–84.

Park, D. C., Morrell, R. W., & Shifren, K. (Eds.). (1999). *Processing of medical information in aging patients: Cognition and human factors perspectives* Mahwah, NJ: Lawrence Erlbaum.

Seckler, A. B., Meier, D. E., Mulvihill, M., & Cammer Paris, B. E. (1991). Substituted judgment: How accurate are proxy predictions?. *Annals of Internal Medicine*, *115*, 92–98.

Smyer, M., Schaie, K. W., & Kapp, M. B. (Eds.). (1996). *Older adults' decision making and the law* New York: Springer.

Zimny, G. H., & Grossberg, G. T. (1998). *Guardianship of the elderly*. New York: Springer.

Aging and Culture

Toni C. Antonucci and Hiroko Akiyama

University of Michigan, Ann Arbor, Michigan, USA

GLOSSARY

aging Becoming old and showing the effects or the characteristics of increasing age.

culture The totality of socially transmitted behavior patterns, norms, beliefs, institutions, and all other products of human work and thought.

demographics The statistical characteristics of human population.

families All members of a biologically or legally related group usually living in one household.

health The condition of being sound in body, mind, and/or spirit; it especially refers to freedom from physical disease and pain.

immigration Entering and settling in a country or region to which one is not native.

life course The interval of time between birth and death.

Although the world population has been aging at an unprecedented rate, the speed of change is considerably different depending on national history and life circumstances. The association between age and culture can be best understood in the light of the life span perspective and the family. Different cultural backgrounds affect individuals' aging experiences contemporaneously and cumulatively. Family is both the conveyer and interpreter of culture. The life span attachment model and the convoy model emphasize that close relationships accumulate over time and differentially affect the aging process and, consequently, the physical and psychological well-being of individuals who live in different cultural milieus.

1. INTRODUCTION

The study of aging has a long and ancient history, with notable comments having come from historical luminaries such as Aristotle, Socrates, and Plato. Despite the increasingly youth-oriented society, the modern study of aging has garnered increased attention during recent years for many reasons, not the least of which is recognition of the significantly changing world demography. This article first presents a brief overview of these changes. It then considers the association between aging and culture, noting the role of family and cultural diversity as well as the implications of both family and culture for health and well-being among the elderly. The article closes with a consideration

of how culture shapes aging and health in the United States as well as around the world.

2. AGING AND DEMOGRAPHICS

Changing world demographics are nowhere as dramatic as they are with respect to the age structure of the world population. According to the UN Populations Division and the U.S. Census, the world is graying at an unprecedented rate. For example, comparison of the percentages of people who will be age 60 years or over in the years 2005 and 2030 indicates that approximately 20% of the European population will be at least 60 years old in 2005 and that 30% of the population will share that status by 2030. The situation in the United States is not quite as dramatic, but the increase is still noteworthy. In 2005, 17% of the U.S. population will be age 60 years or over, and in the year 2030, one-quarter of the U.S. population will be in that age bracket. The figures are much smaller for Latin America, where 9 and 15% of the population are expected to be at least 60 years old in 2005 and 2030, respectively. The lowest figures are for Africa, where only 5% of the population is currently age 60 years or over and a mere 7% is expected to reach that age bracket by 2030. The AIDS epidemic is actually reducing the life expectancy in some African countries. These differences make it clear that the age distribution of the world is changing differently, depending on national history and current life circumstances. Interestingly, and perhaps contrary to the assumptions of most people, the United States ranks only 28th in "aging" among the oldest countries of the world, with most European countries and Japan reporting older populations. Although these changing age structures reflect increased longevity, this longevity is due mostly to changes at the other end of the life span. That is, fewer children are dying in childbirth, thereby increasing the numbers of people who will live to old age, and fewer children are being born, thereby reducing the proportion of a population of young people and naturally increasing the proportion of older people in each society. The decline in fertility has reached critical proportions in many countries, including Italy, Greece, Sweden, Japan, and Spain. Although fertility has been declining in the United States, it has not reached the crisis levels being experienced in other parts of the world.

3. AGING, CULTURE, AND FAMILIES

The effort to understand the association between age and culture must include consideration of both the life span perspective and the family. The life span perspective incorporates a long-term view of the growth and maturity of individuals throughout their lifetimes on multiple dimensions. Development on these dimensions is influenced by previous developmental periods that can manifest through stability or change as well as through continuity or discontinuity. Consequently, individuals become more different from each other as they develop over time. This is noteworthy because life span development both affects and is affected by the cultural milieus within which individuals live. During recent times, those cultural milieus have been changing in unprecedented ways in the United States as well as around the world. Examples of cultural milieus include specific individual experiences (e.g., ethnic identity) as well as more universal experiences (e.g., the unprecedented advances in technology experienced during recent years).

Although the life span perspective focuses on the individual, any effort to understand the aging experience must also include the families within which most people age. The family usually surrounds an individual with a protective convoy that nourishes and supports the individual as he or she faces the challenges of life. The family also is often the embodiment of the cultural perspective that the individual values. Just as each individual emerges as an older person who has evolved throughout his or her lifetime, so too does the individual bring that lifetime of growth and development to the family. At the same time, the family also evolves. A family might begin as a young couple who are the parents of a small child and evolve to a family composed of an adult child with aging parents and later to a family that includes grandparents with adult children and grandchildren. The family can be thought of as the vessel within which the individual experiences culture. Although people experience culture at multiple levels—as an individual, as a family member, and as a member of society—it is the family that most often is both the conveyor and interpreter of the culture. The United States may not be the oldest country in today's industrialized world, but the increased longevity of Americans means that most individuals will at some point be members of three-, four-, or even five-generation families. Hence, the role of the older person as a

family member in both promoting and maintaining culture has been steadily increasing.

4. CULTURAL DIVERSITY AND AGING

The United States has, at least ostensibly, always embraced a commitment to the blending of many cultures. This phenomenon is perhaps best exemplified in the term "melting pot." This term was used to refer to the blending of many cultures, but these cultures were mostly Western European. Although immigrants also came from China (e.g., to work on the railroads) and from Africa (e.g., to work as slaves), those immigrants were generally not among the "blended." An important new development in the cultural influences on aging in the United States is the recent immigration from different parts of the world, namely, Central and Latin America, the Middle East, and Southeast Asia. The nature of this immigration is important not only because it represents different cultural influences but also because the immigration patterns are affecting how the United States ages. As noted earlier, the country is aging rapidly, but particularly important is the fact that among certain ethnic and racial minority groups (i.e., the "new" immigrant groups), aging is taking place at an even more rapid pace because more older people are immigrating to be with their children. Hence, as the U.S. population ages, it is also becoming more diverse both racially and ethnically. This will pose a challenge as different immigrant groups seek to maintain their identities while at the same time embracing the American culture.

Several kinds of aging experiences that will be directly influenced by cultural background and perspective can be identified. For example, some older people came to the United States as young immigrants and are now aging. Having arrived during childhood or early adulthood, they are familiar with the American culture and understand it, often while maintaining some "old world" values. At the same time, their children and grandchildren were most likely born in the United States and have embraced only American values. Some problems may arise if, as these individuals age, they identify more strongly with their cultures of origin, especially in terms of their expectations about the role of families and the experience of aging.

Potentially even more problematic is the phenomenon of immigrating in old age. Some people came to the United States to join their children, who had immigrated earlier and settled in the country. The motivation could be emotional ties (i.e., a desire to be close to their children as they age), or it could be based on pragmatic need (i.e., the expectation that children will care for their aging parents). The potential for conflict in such families is even greater because the elderly immigrants are less likely to have embraced American values; therefore, the situation may exist where parents and their adult children have fundamentally different expectations about the aging experience.

It is clear that people, whether they recognize it or not, experience age through a cultural lens. Thus, the changing racial and ethnic population of the United States represents a unique group of individuals. These different cultural backgrounds affected their individual lifetime experiences as children, adolescents, and adults. It will, of course, also affect their aging experiences both cumulatively and contemporaneously. Each cultural group has a unique history. For example, consider the case of Cuban immigrants in Miami, Florida, who fled Fidel Castro's dictatorship and arrived in the United States feeling forcibly expelled from their homeland. When they arrived in the country, they found very sympathetic Americans who believed that Castro was an evil and unjust dictator who was treating Cuban people unfairly. On the other hand, Middle Easterners also have come to the United States to escape an unpopular dictator, but during recent times, Americans have become more hostile, rather than less hostile, toward people from the Middle East. In addition, despite the fact that Middle Easterners also came to escape an unjust dictator, war, and oppression, they are currently viewed with suspicion, unkindness, and resentment. But another current example indicates that these more recent reactions are not ubiquitous. Interestingly, although the United States is experiencing increased hostilities with North Korea, this hostility does not seem to extend to Koreans in this country. They also fled a homeland torn by war, but somehow Americans are more accepting of their presence and their culture. The attitudes of the American public certainly influence the aging experience of both early and later immigrants.

Although these comments have addressed the general concept of culture, there are culture-specific expectations and beliefs concerning health and aging. The next section briefly considers the health implications of culture on aging persons and their family.

5. AGE, CULTURE, AND SOCIAL RELATIONS

Individuals age within families, communities, and society. They interact with these organized structures most frequently through social relations. Social scientists have studied social relations extensively, and it is increasingly understood that social relations affect both the psychological health and physical health of people of all ages. This is evident in models emerging from the childhood literature, such as the life span attachment model, as well as in those emerging from the adult literature, such as the convoy model of social relations. Both of these models emphasize that social relations accumulate over time and influence both contemporaneous and longitudinal choices. These literatures have identified several elements of social relations that influence both morbidity and mortality. In the convoy model, as well as in other adult social relations models, these elements include structural characteristics such as the number, age, gender, role relationship, and years known of network members; support exchanged with network members such as love, tangible aid, and advice; and evaluation of these relationships, that is, whether one feels that they are positive or negative and whether one is satisfied and/or happy with them. Although some elements of social relations transcend age and culture, other aspects of social relations are fundamentally influenced by culture and age. For example, most people of all ages have close social relations with their parents, spouses, and children. Europeans tend to include more extended family members among their close social relations, whereas this is less true among Americans, who are more likely to include friends.

Social relations involve the exchange of social support in all cultures. How often these supports are exchanged, as well as who exchanges them, what supports are exchanged, and when the supports are exchanged, often varies by age and cultural expectations. Although it is common across cultures to expect that parents will provide for their children when their children are young, the degree to which parents continue to do so when those children become adults, or the way in which the expectation might reverse such that parents expect support from their adult children, varies by culture and country. In a recent Italian court case, parents were sued by their employed, 35-year-old son because he was reluctant to accede to his parents' wishes that he move out and live on his own. The

Italian court ruled in favor of the son, requiring the parents to continue providing room and board for as long as the son so desired. A similar case would be unheard of in the United States because it is non-normative for an able adult son to live with his parents and because no court would sustain such a petition. Similarly, in the United States, it is common for parents to attempt to provide for their adult children, for example, through gifts, favors, and loans. But in Japan, the culture assumes the opposite; that is, no matter what age, children are understood to be in a state of indebtedness to their parents and, thus, are expected to be constantly in the position of providing to their parents rather than receiving from them.

There are many factors that influence how social relations are evaluated. There are objective characteristics that influence this evaluation, but a solid accumulation of empirical evidence indicates that significant subjective characteristics influence this evaluation as well. Subjective evaluation could be considerably different depending on the age and culture within which the social relations are embedded. The Italian son just described would rate his parental support received as poor, whereas an American son would not. An adult American son might be dissatisfied and unhappy if his parents did not provide a college fund or a downpayment on a house, whereas the lack of either of these in another culture would generate no parallel dissatisfaction or unhappiness. For older family members, it is clear that their expectations vary depending on both age and culture. Interestingly, in the United States, some evidence indicates that younger people expect to provide more support than their elders expect to receive, whereas in other countries, parents report rather high expectations of what their children should provide. Regardless of the cultural convention regarding who is in an individual's social network or an objective assessment of what is exchanged, evaluation of social relations varies considerably. This variation, although influenced by many factors, has been well documented in the influence of culture on health and well-being.

6. CULTURAL INFLUENCES ON THE HEALTH OF AGING PEOPLE

Different cultural experiences, traditions, and values have direct implications for the expectations of both elders and their families. These expectations range from minor

considerations (e.g., concerning customs around favorite foods or holiday traditions) to more significant ones (e.g., concerning family roles and responsibilities).

There is clear evidence that culture influences health. Some examples are illustrative. It has been documented that elderly Japanese living in Japan are much more likely to die of a stroke than are elderly Japanese living in the United States, where the leading cause of death has been cardiovascular disease. This difference has been traced to dietary differences in ethnically Japanese people living in Japan versus those living in the United States. Similarly, hypertension is a major health problem among people of African descent living in the United States but is relatively rare among Africans living in Africa. Researchers have suggested that this difference is caused by the stressful life circumstances of racial discrimination experienced in the United States. Another interesting example is menopause. Research in medicine, epidemiology, and anthropology has consistently documented that women in many countries (e.g., Japan) experience fewer and/or lighter menopause symptoms (e.g., hot flash, insomnia, depression) than do American women. Although a consensus explanation has not yet been reached, these differences have been attributed to diet (e.g., soybeans) and cultural attitude toward (e.g., acceptance of) the physical and physiological changes associated with aging. A final and compelling argument for the influence of culture on health is that, regardless of the health differences evident among people of different countries, by the third generation all immigrants exhibit patterns more similar to other Americans than to people from their countries of origin.

Cultural expectations of family roles and responsibilities are critical when they have a direct impact on an older person's health and expectations concerning health care. Some simple examples are illuminating. In some countries (e.g., Japan), it is much more common for older people to live with their adult children. The differences in customs have both positive and negative implications. In Japan, older people often live in their own home, and an adult child—often the eldest son—continues to live at home with his wife and children. Although this is a traditional custom that has been changing in modern Japan, it does mean that traditionally a live-in relative attended to the health care needs of a family's oldest members. This live-in relative was most often not the eldest son but rather his wife, that is, the daughter-in-law. Much has been made of the advantages of coresidence for the elderly in terms of meeting their health care and other needs as they get older. However, social scientists have increasingly documented that not all such arrangements are ideal. There is often tension between the daughter-in-law and her parents-in-law as well as between the son and his parents. This coresidential custom is much less common in the United States, and when it does occur, it is often the case that a parent moves in with an adult child, although the opposite may be the case when the adult child is impoverished or experiencing a life crisis such as divorce or job loss. Under the coresidential condition where the elder moves in with a child, the power dynamics are reversed in terms of the traditional American family dynamics as well as the situation just described in Japan. Tension can arise in this case as well, although health care needs and expectations are also more likely to be met. Certainly, these examples provide persuasive testimony concerning the influence of culture on the health experience of aging people.

7. SUMMARY AND CONCLUSIONS

With the radically changing world demographics leading to a significant increase in the number of older people, it is critical to attend to those factors that positively or negatively affect the experience of aging. Although culture is an important influence, one should be aware that there are also other factors that influence aging and family relations such as socioeconomic status, race, and industrialization. Therefore, culture should be neither overlooked nor overestimated. It remains for researchers and clinicians to identify when culture is a primary factor and when it is not and to determine at what point one should seek to optimize the experience of aging both within and across cultures.

See Also the Following Articles

Age-Related Issues among Minority Populations ■ Aging and Competency ■ Anxiety Disorders in Late Life ■ Cognitive Aging ■ Cognitivism ■ Depression in Late Life ■ Elder Caregiving ■ End of Life Issues ■ Psychotherapy in Older Adults

Further Reading

Antonucci, T. C. (2001). Social relations: An examination of social networks, social support, and sense of control. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 427–453). San Diego: Academic Press.

Blieszner, R., & Bedford, V. (1995). *Handbook on aging and the family*. Westport, CT: Greenwood.

Jackson, J. S. (Ed.). (1998). *The Black American elderly: Research on physical and psychosocial health*. New York: Springer.

Jackson, J. S., Antonucci, T. C., & Gibson, R. C. (1995). Ethnic and cultural factors in research on aging and mental health: A life-course perspective. In D. K. Padgett (Ed.), *Handbook on ethnicity, aging, and mental health* (pp. 22–46). New York: Greenwood/Praeger.

Keith, J., Fry, C. L., Glascock, A. P., Ikels, C., Dickerson-Putman, J., Harpending, H. C., & Draper, P. (1994). *The aging experience: Diversity and commonality across cultures*. Thousand Oaks, CA: Sage.

Sokolovsky, J. (Ed.). (1997). *The cultural context of aging*. Westport, CT: Bergin & Garvey.

Aging, Cognition, and Medication Adherence

Odette N. Gould

Mount Allison University, Sackville, New Brunswick, Canada

GLOSSARY

external memory strategies Actions undertaken to support prospective memory whereby the physical environment is changed (e.g., the use of lists, notes, calendars, and pillboxes).

internal memory strategies Mental actions undertaken to support prospective memory (e.g., linking planned activity with daily events).

medication nonadherence Failure to take medications as prescribed; this can include taking too few or too many doses, taking medication at incorrect times or in incorrect quantities, or not following special instructions related to the medication (e.g., regarding food or beverages).

prospective memory Memory for actions to be performed in the future (e.g., remembering to take prescriptions at a prescribed time).

reality monitoring The ability to distinguish between actions that were planned and actions that were actually carried out.

retrospective memory Memory for events that occurred in the past (e.g., remembering the instructions received for when and how to take medication).

Medication adherence is increasingly recognized as a complex task involving many cognitive abilities, and adherence errors are common in both younger and older adults. Patients must integrate often complicated instructions on how to take medication into their daily routines, remember to take the medication at the appropriate time, and remember that the medication has been taken. Although age-related declines in many cognitive abilities are clear, older adults manage familiar tasks such as medication taking quite well, at least until very late age. Many mnemonic supports have been found to function quite effectively to remind patients to take their medication. However, more and more, researchers and health care providers are focusing on multifaceted interventions that are tailored to the specific needs, abilities, and beliefs of older patients.

1. EXTENT AND COST OF MEDICATION NONADHERENCE IN THE ELDERLY

Individuals over 65 years of age make up an increasingly large percentage of the population in industrialized

countries around the world, and the trend is expected to continue for many decades to come. Indeed, the World Health Organization has projected that the percentage of older adults in developed countries will increase from 14.3% in 2000 to 19.2% by 2020. The majority of older adults suffer from at least one chronic illness; consequently, older adults are the largest consumers of medications of any age group. Many of the illnesses that afflict older adults can be treated or at least controlled quite successfully through the use of prescription medication—if medications are taken as prescribed. Unfortunately, rates of adherence to medication regimens are believed to be quite low. It is quite difficult to establish adherence rates precisely, but in general, as few as 50% of patients treated for chronic conditions are believed to take their medications as prescribed. Rates of adherence are also thought to be particularly low in asymptomatic conditions such as hypertension that afflict many older adults. Surprisingly, most studies have not found evidence that older adults in general are more nonadherent to medication than are middle-aged or younger adults. Recent work by Park and colleagues even indicated that older adults, between 65 and 74 years of age, made the fewest number of adherence errors of any age group, although adults over 75 years of age made the most.

The risks of nonadherence are numerous and serious. Up to 40% of hospital admissions in adults over 65 years of age are due directly to medication errors. Even though older adults do not seem to be more nonadherent than younger adults, they are more likely to experience serious health consequences because their bodies are less tolerant of medication errors and drug interactions. When treatment regimens are not followed, the risk of toxicity is heightened, medical conditions are likely to worsen, and expensive and often invasive treatments become necessary. Clearly, enhancing medication adherence is a priority for the effective treatment of chronic diseases, for enhancing the well-being of elderly individuals, and for reducing health care costs.

2. TYPES OF NONADHERENCE

Medical nonadherence occurs when patients do not carry out medical regimens as prescribed and includes any of the following: (a) not taking the medication at all, (b) taking smaller or larger doses than prescribed, (c) taking fewer or more doses than prescribed, or (d) not following recommendations regarding food or beverages while medicated. Such behaviors can be either intentional (i.e., the patient chooses not to take

the medication as prescribed) or unintentional (i.e., the patient intends to take the medication correctly but does not). In general, unintentional overuse or underuse of medication is due to cognitive variables such as forgetfulness and misinterpretation of instructions. Intentional noncompliance, on the other hand, can be seen as relatively more complex and is linked to either external circumstances or motivational issues. External circumstances include factors such as not being able to pay for the medication, not being able to reach the pharmacy to get the prescription filled, and not being able to open resistant packaging (e.g., childproof bottle tops) due to physical handicaps. Motivational issues refer to the fact that many individuals, both young and old, hold views about their illness and their medication that are incompatible with adherence.

3. ADHERENCE: A COMPLEX COGNITIVE TASK

In the past, medication adherence has been presented as an example of a simple everyday prospective memory task, where prospective memory is defined as remembering to do things in the future. Recently, however, there is a clear consensus emerging that medication adherence is a complex process involving a host of variables. Although many of these factors are clearly cognitive, others (e.g., motivation to adhere to the medication regimen) involve the interplay between cognitive and social factors.

From a cognitive standpoint, the patient must understand the instructions for how to take the medication. These instructions are often complex, especially when multiple medications, each with its own regimen of doses and times, are presented. Thus, combining the information across multiple medications, and integrating all of this information with daily activities to form a plan of action, can be quite challenging. Once the plan of action is created, the patient must remember to take the medication at the correct time and must remember what dose and what medication to take. Moreover, the patient must keep track of whether or not the medication was taken so that extra doses are not taken inadvertently (i.e., reality monitoring). Reality monitoring can be particularly challenging when a medication for a chronic condition has been taken at the same time of day for months or years. Finally, the patient should monitor his or her health for a range of possible side effects from the medication, including symptoms such as confusion and forgetfulness.

4. PROSPECTIVE MEMORY AND EPISODIC MEMORY: AGE-RELATED CHANGES AND EFFECTS ON ADHERENCE

Clearly, cognitive factors such as language comprehension (e.g., understanding the medication instructions), long-term memory (e.g., remembering what to do), working memory (e.g., juggling the competing demands of everyday tasks and medication taking), problem solving (e.g., integrating complex medication instructions with the daily routine), prospective memory (e.g., remembering to take the medication at the correct time), and metamemory (e.g., monitoring memory performance and judging whether external support is needed for adequate adherence) all are involved in determining whether correct adherence will occur. Although age-related losses in many of these areas have been clearly established, many older adults are able to maintain high levels of performance, especially on everyday tasks such as medication taking.

Many language abilities are maintained into very old age. However, the comprehension of complex medication instructions may depend heavily on working memory capacity as well as language abilities. Age-related losses in working memory and problem solving are well documented. It is clear that if information is presented quickly or is presented using an unfamiliar or disorganized format, many older adults will have difficulty in comprehending and recalling the information. Many studies have shown that older adults are at a distinct disadvantage, compared with younger adults, when asked to combine complex medication instructions into a plan of action, especially if the instructions require inferences. Such problems have also been found when older adults are asked to transfer medication into weekly pillboxes, especially if the pillboxes do not contain multiple compartments for each day.

To take medications correctly, the patient not only must understand the instructions when they are given but also must remember these instructions over the days and weeks following the meeting with the physician. In general, age-related losses are also observed when the ability to recall verbal information after a relatively long interval (i.e., hours or days) is measured. However, age differences can be greatly reduced, at least for relatively healthy older adults, if environmental supports for recall are present. For example, if information is presented clearly, if multiple formats are used (e.g., written and verbal presentation), if recall is based on habitual or familiar activities, if cues for recall are present in the environment, and if the person is not distracted by competing stimuli at time of recall, age differences in episodic memory are greatly reduced or even eliminated.

Although it is well documented that there are age-related losses in working memory, problem solving, and long-term memory, prospective memory does not show the same rates of decline with age. For example, when asked to make a telephone call or mail a postcard at a certain time, older adults attain performance levels similar or superior to those of younger adults through the use of memory strategies. This may help to explain why older adults do not make more medication errors than do younger adults. Park and colleagues have speculated that adults in their 60s and 70s may have few adherence errors because (a) they are sufficiently focused on their health to adopt mnemonic strategies for medication (unlike middle-aged adults) and (b) they have the cognitive ability to adopt and use mnemonic strategies (unlike much older adults).

In laboratory-based studies of prospective memory, it has been suggested that prospective memory tasks can be divided into time-based and event-based tasks. In a time-based task, the individual must accomplish the task at a certain time (e.g., "Take your pill at 8 AM"), whereas in an event-based task, the individual must accomplish the task when cued by a specific event (e.g., "Take your pill before breakfast"). Computer-based simulations of prospective memory tasks have shown that event-based tasks are easier for older adults to accomplish, especially if the accompanying background task is not particularly difficult. This may also explain why many older adults report that the main strategy they use to remember their medication is to mentally link the medication with routine daily activities. It may also explain why some studies have found that tailoring medication regimens to patients' routines enhances adherence significantly.

5. KNOWLEDGE, BELIEFS, AND COMMUNICATION: AGE-RELATED CHANGES AND EFFECTS ON ADHERENCE

At least three major foci in the research on the motivational factors of adherence in old age can be identified. First, the beliefs and attitudes that the patient has about his or her illness have been shown to predict adherence. Some people may deny that they are ill or might not believe that their conditions have serious consequences.

Patients who are more knowledgeable about their conditions are more likely to follow recommended medication regimens. This may also explain why social support has been found to predict adherence. It may be that significant others play a role in motivating an individual to adhere. Providing cues or reminders to follow the prescription also may play a role.

A second factor concerns the beliefs that the patient holds about his or her treatment. Beliefs about the effectiveness of the medication, about whether the regimen prescribed is appropriate, about whether medications cause uncomfortable side effects, and about how important it is to adhere to medication regimens as prescribed have been shown to contribute to treatment adherence. Thus, some older adults may simply not believe that their medication is helpful. Others may choose underadherence to avoid side effects or may overmedicate because they believe that the prescribed dose is ineffective. (It should be noted that because many health care professionals lack extensive knowledge about the effects of medications—especially multiple medications—on the older body, partial adherence may be beneficial in at least some cases.)

Finally, a third set of factors revolves around the relationship with the health care provider. This relationship is critical in many ways. Medication adherence can be seen as the outcome of a process that begins before the physician and patient even meet. Namely, the attitudes and beliefs that each holds about the other are likely to affect patient–physician interactions in at least two ways. First, the quality of the interaction and the satisfaction that the patient feels with his or her health care are very important predictors of motivation to adhere to a medication regimen. Second, the relationship may also affect the quantity and quality of the information that is exchanged during the patient–physician encounter. For adherence to occur, the patient must be told how to take the medication in detail and must be informed about his or her condition and its associated risks. Unfortunately, many studies suggest that health care providers are less likely to provide this information to older patients than to younger ones, and older patients are also less likely to question their physicians. Thus, even though older adults are likely to be taking more medications and to have more complex regimens, they are also likely to receive less information about their medications. Finally, the relationship between the health care provider and the patient also affects the quality of the information that the patient provides. In other words, the patient is more likely to provide accurate and detailed information about his or her symptoms and adherence to existing medications when the patient perceives the health care provider to be caring and approachable. Such accurate information is key for the health care provider to be able to make accurate diagnoses and treatment decisions.

In conclusion, even though the relationship between the health care provider and the patient is commonly addressed from a social psychology perspective, this relationship has important cognitive repercussions as well. The patient's knowledge about the illness and medication and his or her beliefs about the importance of adherence are often highly dependent on how well and how comfortably the patient communicates with the health care provider.

6. MEASURING ADHERENCE: A DIFFICULT CHALLENGE

The difficulty of measuring medication adherence is well established and often discussed in the literature. Both technical and nontechnical approaches have been used. One technical approach has been the evaluation of body fluids (e.g., blood, urine) for the presence of the prescribed medication. However, this technique only recognizes drugs ingested during the recent past and does not indicate whether the patient has taken the medication regularly. Pharmacy refill rates and pill counts have also been used; however, with these methods, flawed data are obtained if the patient is using multiple pharmacies, disposing of pills, hoarding pills, and/or sharing pills with others. The most accurate data come from a specially designed pill bottle with a lid containing a microchip that records every time the bottle is opened. However, even this system still does not provide an accurate measure of whether the drugs were actually taken when the pill bottle was opened and does not verify whether the correct doses were taken. Although these special pill bottles are useful for research purposes, their general use remains limited because they are very costly.

The most convenient and the most often used means of assessing adherence is self-reports. However, it is clear that self-reports underestimate medication errors. Patients are likely to overestimate their adherence levels, not only because they do not recall errors but also because they hope to avoid confrontations and/or embarrassment. Cognitive factors also contribute to the accuracy of self-reports. When asked to recall the frequency of a behavior, respondents are likely to reduce their efforts

by estimating rather than counting, and this is especially true of older adults. Moreover, with older adults, memory complaints are often not predictive of objective memory performance. Thus, the older adults who complain about having bad memories might not be the ones who make adherence errors due to forgetting. Despite the potential pitfalls of self-reports, recent work by Hertzog and colleagues suggests that when older patients are asked specific questions about memory problems relating to medication adherence rather than general memory performance, self-reports can be quite useful.

7. INTERVENTIONS: HOW COGNITIVE PSYCHOLOGY CAN HELP

Much of the intervention research has focused exclusively on compensating for either cognitive or motivational deficits. Many studies have focused on providing older adults with external mnemonic strategies. Such strategies involve effecting changes in the environment so that it provides prospective memory cues. Many studies have been carried out to investigate the effectiveness of aids such as pillboxes and pill bottle alarms, memory training, voice mail reminders, and organizational charts. Most of these studies found such memory aids to be at least somewhat effective in enhancing medication adherence. However, many of these studies were conducted over a short period of time, and neither young nor older adults are likely to continue using memory strategies assigned to them after the training period has ended. It is also interesting to note that the few studies that examine what strategies older adults use spontaneously tend to find that internal strategies (i.e., using mental actions to help encode or retrieve information) are reported quite often. However, very little is known about the effectiveness of internal strategies for routinized, long-term prospective tasks such as medication adherence. Also, little is known about what drives the process of determining that a mnemonic strategy is necessary, deciding which one to use, and evaluating its effectiveness.

One external strategy that is being adopted by many older adults is the use of pillboxes with daily compartments. However, as noted previously, at least some older adults have difficulty in transferring pills to the container correctly, at least when the smaller versions are used (i.e., only one compartment per day), and the larger containers have the disadvantage of being bulky and might not be practical unless patients are

housebound. Recently, attempts to reduce adherence errors have focused on using technological advances in how medication is presented to patients. Some have speculated that as pharmaceutical science continues to advance, medications will be released into the body in a controlled fashion (e.g., slow-release pills), thereby making frequent doses and complex regimens unnecessary. More immediately, the use of individualized packaging is becoming more affordable and common. For example, when blister packs are used, the time to take the medication can be listed on the package, and monitoring whether the medication has been taken becomes much easier. However, when patients must take many different medications, each in its own blister pack, the packaging may increase rather than decrease the confusion (and be large and bulky to transport as well). Finally, increases in adherence levels have also been obtained in studies that addressed motivational factors. These studies focused on improving the relationship between the physician and the patient or tried to educate the patient as to the seriousness of the illness or the importance of the medication.

In conclusion, many different interventions can be useful, but there is a growing recognition that the intervention applied must be tailored to the characteristics of the patient. Clearly, it is not helpful to provide mnemonic strategies to someone who is nonadherent by choice, and neither is it helpful to convince someone of the importance of medications if the person cannot remember to take them. Thus, cognitive psychologists today are working in collaboration with health care providers to develop exciting multifaceted approaches that tailor cognitive and motivational interventions to the needs, abilities, and beliefs of individual older adults.

See Also the Following Articles

Age-Related Issues among Minority Populations ■ Aging and Competency ■ Cognitive and Behavioral Interventions for Persons with Dementia ■ Cognitive Skills: Training, Maintenance, and Daily Usage ■ Elder Caregiving ■ Psychotherapy in Older Adults

Further Reading

Brandimonte, M., Einstein, G. O., & McDaniel, M. A. (1996). *Prospective memory: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum.

Johnson, M. J. (2002). The medication adherence model: A guide for assessing medication taking. *Research and Theory for Nursing Practice: An International Journal, 16*, 179–192.

McElnay, J. C., & McCallion, R. (1998). Adherence in the elderly. In L. B. Myers, & K. Midence (Eds.), *Adherence to treatment in medical conditions* (pp. 223–253). Amsterdam: Harwood Academic.

Murdaugh, C. L. (1998). Problems with adherence in the elderly. In S. A. Shumaker, E. B. Shron, J. Ockene, & W. L. McBee (Eds.), *The handbook of health behavior change* (2nd ed., pp. 357–376). New York: Springer.

Park, D. C., & Jones, T. R. (1997). Medication adherence and aging. In A. D. Fisk, & W. A. Rogers (Eds.), *The handbook of human factors and the older adult* (pp. 257–287). San Diego: Academic Press.

Park, D. C., Morrell, R. W., & Shifren, K. (1999). *Processing of medical information in aging patients: Cognitive and human factors perspectives.* Mahwah, NJ: Lawrence Erlbaum.

Agreeableness

Brad E. Sheese and William G. Graziano

Purdue University, West Lafayette, Indiana, USA

GLOSSARY

effortful control The ability to suppress a dominant response for a subdominant response (e.g., the ability to cope with frustration).

knowledgeable others Peers, teachers, parents, and others who know the person being evaluated.

latent variable An underlying system.

personality dimension A set of related traits.

Agreeableness is a personality dimension (or set of related traits) that describes a class of individual differences that generally have to do with being pleasant, likable, and harmonious in relations with others. Empirical research finds that Agreeableness as a dimension hangs together consistently (persons who are "kind" also tend to be "considerate"), is relatively stable and enduring over time, and is related to many kinds of human activity.

1. DESCRIBING AGREEABLENESS: WHAT IS IT?

Agreeableness has become the label most frequently used for this personality dimension, but it is only one of many such labels. Some of the other labels used to describe the dimension (or closely related dimensions) are tendermindedness, friendly compliance versus hostile noncompliance, love versus hate, likability, communion, and conformity. It has been argued that none of these labels, including Agreeableness, adequately captures either the breadth or the substantive content of this dimension of personality. As a label, Agreeableness has been criticized specifically for being too narrow and perhaps for overemphasizing acquiescence. Theorists have suggested that it may be more appropriate to refer to the dimension either with numerals (the Roman numeral II has been used in the past) or simply with the letter A (for agreeableness, altruism, and affection).

At a theoretical level, Agreeableness describes an underlying system (latent variable) of individual differences. It is one of five broad personality dimensions that appear in all versions of the five-factor approach to personality (i.e., the Five-Factor Model). The five-factor approach describes personality at perhaps its broadest and most abstract (decontextualized) level. Trait adjectives that are positively associated with Agreeableness include kind, warm, cooperative, unselfish, polite, trustful, generous, flexible, considerate, and agreeable. Trait adjectives that are negatively associated with

Agreeableness include cold, unkind, uncooperative, selfish, rude, distrustful, stingy, stubborn, and inconsiderate. Overall, Agreeableness describes a broad, but related, set of individual differences in how a person relates to others. Specifically, Agreeableness appears to describe differences in being predominantly prosocial or other-oriented versus antisocial or self-oriented in social interactions.

Self-report measures of Agreeableness are those most frequently used for assessment. The use of ratings of Agreeableness by knowledgeable others (e.g., peers, teachers, parents) in addition to self-ratings is desirable from a psychometric standpoint but is less common due to the increased demands associated with collecting this information. Currently, there are several measures available to assess Agreeableness within the framework of the five-factor approach to personality. The three most frequently used measures in research are Goldberg's adjective markers, the questionnaire-format Big Five Inventory (BFI) developed by John, and the Neuroticism Extraversion Openness Personality Inventory–Revised (NEO) developed by McCrae and Costa. These measures vary in their construction, length, and content, yet empirical research shows that the five-factor dimensions assessed by these measures, including Agreeableness, are nearly identical across measures. Focusing on Agreeableness, there is some discrepancy between the NEO and the other two measures. The main difference is that the NEO includes warmth as a facet of Extraversion, whereas in the other two measures warmth is associated with Agreeableness. Overall, however, outcomes of studies using different measures of Agreeableness show convergence. For most purposes, the various measures of Agreeableness are functionally equivalent.

2. EXTERNAL CORRELATES: WHAT IS AGREEABLENESS RELATED TO?

Over the past decade or so, cross-cultural research has uncovered the same or similar five-factor personality structure in many cultures with many different languages and in both adults and children. In all cultures studied, Agreeableness has been seen as a major dimension of personality. At a minimum, this research indicates that Agreeableness describes a set of individual differences that are of interest worldwide. The dimension is not an artifact of the English language, college-age participants, or Western cultural norms. Recent research also suggests that despite the clear preference that most groups have for persons with agreeable qualities over persons with disagreeable qualities, Agreeableness differences are also not an artifact of social desirability, self-presentation, or self-deception, nor are they merely a characteristic of observers' evaluations. Instead, research suggests that Agreeableness describes a stable enduring set of qualities in individuals that is related to important social outcomes.

Agreeableness differences can be studied in many different ways. One of the most common ways is to bring persons who differ on Agreeableness into a laboratory, where their behaviors can be assessed in various kinds of social situations. Typically, these persons are college students. In laboratory studies of group processes, Agreeableness is related to lower within-group conflict and to positive within-group evaluations. For example, more agreeable people like their fellow group members more and are liked more in return. Research also shows that Agreeableness is negatively related to competitiveness in group interactions and is positively related to expectations of enjoying group interactions. These findings suggest that Agreeableness may contribute to group cohesion and may serve as a protective factor against group dissolution. These findings also provide empirical support for the notion that Agreeableness can be mapped onto the social motives described by communion.

In some circumstances, being a more cooperative and generally more prosocial group member may contribute positively to group and individual functioning. In these circumstances, one would expect that Agreeableness is related to both group- and individual-level performance. In some situations, however, being considerate, trusting, and compassionate may undermine a group's performance or an individual's performance. In these situations, Agreeableness may be unrelated, or even negatively related, to performance. One meta-analysis found that Agreeableness was negatively related to job performance for managers but also found that Agreeableness was positively related to job performance for professionals and skilled laborers. A later meta-analysis found that in most studies, Agreeableness was unrelated to job performance but was positively related to teamwork. In a competitive business environment, Agreeableness may be a more desirable quality for group members than for managers or business leaders. To a certain extent, goals and tasks that emphasize competition, profits, and self-gain may run contrary to the selflessness and positive social orientation associated with Agreeableness. This may, in part, explain findings that, across the life span, Agreeableness is negatively related

to extrinsic success (e.g., salary) and career satisfaction. Perhaps more agreeable individuals would find more satisfaction, if not more success, in helping professions or public service than in corporate job environments.

Other research (some laboratory and some field studies) has examined Agreeableness and dyadic relationship processes and outcomes. One area of interest has been the resolution of dyadic conflicts. If more agreeable individuals are more motivated to maintain positive social relations with others than are their less agreeable peers, one would expect more positive patterns of attributions and behaviors from agreeable individuals than from their peers when they have disagreements with others. We might also expect that a concern for others would lead to less disagreement or conflict with others in general. Research outcomes have shown that in both adolescents and adults, Agreeableness is positively related to constructive conflict resolution tactics (e.g., negotiation) and is negatively related to destructive conflict resolution tactics (e.g., physical force). Agreeableness is also negatively related to the number of daily conflicts that individuals report in daily diary records. After a conflict has occurred, agreeable individuals say that they are more upset by the conflict but are less likely to have vengeful attitudes about the conflict or to use vengeance as a problem-solving tactic.

In general, more agreeable individuals appear to behave in ways that are constructive rather than destructive to their relationships with others. This suggests that Agreeableness should be positively related to relationship quality (e.g., fewer conflicts, more satisfaction) and perhaps to relationship quantity (e.g., more friends, more robust social networks). Agreeable individuals may incur costs in maintaining their relationships, but they should also gain the benefits associated with having a strong social network. During adolescence, research found that Agreeableness was positively related to the quality of teacher and peer relations as well as to peer social status. In adults, research on social support showed that Agreeableness was positively related to the perceived support satisfaction, and negatively related to the frequency, of negative social exchanges.

These positive relationships with others may either stem from or contribute to a positive relation between Agreeableness and evaluations of the self. As discussed previously, agreeable individuals view others favorably and are viewed favorably by others. Research also shows that agreeable individuals view themselves favorably. During adolescence, Agreeableness and self-esteem are positively related in both self-ratings and teacher ratings. In adults, Agreeableness is positively related to

happiness and subjective well-being. Agreeableness is negatively related to anger, hostility, and (indirectly) depression in both daily and general dispositional self-reports. In general, agreeable individuals earn higher scores on measures of adjustment than do their less agreeable peers. Conversely, research shows that less agreeable individuals are more likely to exhibit various kinds of externalizing problems during both adolescence and adulthood.

3. ETIOLOGY/ANTECEDENTS: WHERE DOES AGREEABLENESS COME FROM?

In considering the development of individual differences in Agreeableness, researchers have taken two distinct but related approaches. The first approach is to examine the specific contribution of genes to individual differences in Agreeableness. Over the years, behavioral genetic studies have attempted to determine the contribution of both genetic and environmental (shared and nonshared) influences on personality. Results have varied, with at least one study suggesting that among the Big Five dimensions, Agreeableness is the most susceptible to environmental influences. However, most of the behavioral genetic studies do not find that Agreeableness is any less heritable than the other Big Five dimensions. Some theorists suggest that all of the Big Five dimensions, including Agreeableness, represent entirely genetically driven, biologically based dispositions. Supporting this claim are a variety of studies showing that constructs related to Agreeableness, such as altruism, prosocial behavior, and empathy, also show moderate to high heritabilities.

The second approach to identifying the root of individual differences in Agreeableness has been to link adult personality to infant and child temperament. One problem is that the individual differences identified through modern temperament approaches rarely map directly onto adult personality traits. Agreeableness, in particular, has no clear antecedent among temperament constructs. This may, in part, have to do with the emphasis of Agreeableness on the quality of peer relations and other-oriented responsiveness. During infancy, it may be difficult to detect individual differences in the concern for the quality of relationships with others given that infants have a comparatively limited repertoire of both behaviors and communication abilities. In addition, in comparison with other personality

dimensions such as Extraversion and Neuroticism, if Agreeableness is already present as a disposition, it may require the development of more elaborate cognitive processes (e.g., perspective taking, theory of mind) before it can be observed.

Among the temperament traits that have been observed, theorists have suggested that individual differences in effortful or executive control may be an antecedent of individual differences in Agreeableness. Effortful control refers to the ability to suppress a dominant response for a subdominant response. In particular, effortful control is related to the ability to cope with frustration and respond adaptively to goal blockage. More generally, researchers have proposed that effortful control is linked to both cognitive (e.g., attention) and emotional self-regulatory capabilities. In this light, individual differences in Agreeableness may reflect differences in the ability to monitor and control a variety of self-centered or even antisocial impulses in social situations. For example, in daily interactions, less agreeable individuals report feeling (and probably exhibit) more anger and hostility than do their more agreeable peers. These feelings of anger and hostility may stem from an underlying inability to regulate or suppress emotions to avoid damaging relationships.

Even if there were large heritable or temperamental contributors to Agreeableness, it would still be necessary to know how social learning histories and life experiences shape these tendencies into adult individual differences. There is good evidence that cultures differ in the values they assign to Agreeableness; accordingly, we might expect average differences in Agreeableness among children reared in different cultures. Nevertheless, differences within cultures remain. Agents of socialization may leave their marks, but persons enter the socialization environment with something more complicated than blank slates. Some individuals will be receptive to agents of socialization, absorbing and internalizing cultural values, whereas other individuals will resist. In this approach, then, socialization can produce large effects on Agreeableness, yet adult differences in Agreeableness within cultures are the end-state products of Person × Socialization interactions.

Consistent with this line of reasoning, recent research has demonstrated complex links between childhood self-regulation and adult Agreeableness. Childhood self-regulation has, in turn, been linked both to the quality of child-rearing practices and to temperamental characteristics. It is possible that temperamental characteristics and socialization experiences may interact in a variety of ways to produce individual differences in Agreeableness. For example, temperamental characteristics may contribute to the selection of particular social ecologies that restrict the range of socialization experiences available to any given individual. Children who lack self-control and are easily frustrated may make for poor relationship partners and may subsequently be rejected by their peers. Consequently, these children might miss important socialization experiences provided by peer relationships. Temperament may interact with parent–child relationships to affect socialization experiences in a similar way. As children get older, they may also consciously choose to place themselves in particular environments that are consonant with their temperamental predispositions. Similarly, as adult personality characteristics emerge, they may interact with situations to reinforce underlying predispositions. For example, if given a choice, highly agreeable children may choose to avoid highly competitive situations. If forced into competitive situations, highly agreeable children may attempt to convert the situation into a more cooperative one or otherwise adapt to the social norms associated with highly competitive interactions.

4. CONCLUSION

Agreeableness is a dimension of personality associated with motives for maintaining positive relations with others. It is a superordinate descriptor for a system of lower order affective, behavioral, and cognitive tendencies. Cultures may differ in the emphasis they put on socializing children into seeking harmonious relations with others, but differences remain within cultures. Agreeableness differences in people seem to be relatively stable over time and across settings. Agreeableness may have developmental origins in temperament and in skills in coping with frustration. These differences are associated with corresponding tendencies in dealing with conflict, emotional responsiveness to others, group interaction, interpersonal relationships, and self-evaluation. For many life events, Agreeableness is associated with positive outcomes such as good social relations, friendships, better adjustment, and social recognition (with the exception of external rewards, e.g., salary). There is some suggestion that agreeable persons are less effective as managers, but there is no evidence (at least so far) for claims that agreeable persons are less attractive, more conforming, more compliant, or more submissive than their peers.

See Also the Following Articles

Emotion ■ Personal Initiative and Innovation ■ Social Networks ■ Traits

Further Reading

Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 795–824). San Diego: Academic Press.

Kohnstamm, G. A., Halverson, C. F., Mervielde, I., & Havill, V. L. (1998). *Parental descriptions of child personality: Developmental antecedents of the Big Five?* Mahwah, NJ: Lawrence Erlbaum.

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement in interpersonal behavior. In D. Cicchetti, & W. Grove (Eds.), *Thinking critically in psychology: Essays in the honor of Paul E. Meehl* (pp. 89–113). New York: Cambridge University Press.

Alcohol Dependence

Norman S. Miller

Michigan State University, East Lansing, Michigan, USA

GLOSSARY

anhedonia Lack of pleasure or satisfaction in experiences that are normally pleasant or satisfying.

delirium tremens A kind of alcohol withdrawal occurring during abstinence or reduced alcohol use after excessive long-term alcohol consumption; symptoms include delirium and hallucinations.

DSM-IV-R Abbreviation for *Diagnostic and Statistical Manual of Mental Disorders,* 4th edition, revised (published January 2000).

DSM-IV-TR Abbreviation for *Diagnostic and Statistical Manual of Mental Disorders,* 4th edition, text revision (published June 2000).

gateway drug A substance that, when abused, is known to lead to the abuse of other substances.

phylogenetic Relating to evolutionary relationships within and between groups.

Practically everyone has consumed alcohol in his or her lifetime. Fully 90% of the U.S. population has had at least one alcohol-containing beverage. In addition, 30% of the population drinks regularly with adverse consequences (15% develops alcohol dependence and another 15% drinks heavily). Costs due to problems from alcohol use and dependence reach $100 billion annually in the United States. Despite the available knowledge about the etiology and course of alcohol consumption, the costs from adverse consequences continue to rise each year for a wide range of age groups (including children, adults, and the elderly), most ethnicities and cultures, and both genders.

1. WHO BECOMES ALCOHOLIC?

The best predictor of who will become an alcoholic is a positive family history of alcoholism. Family history studies of alcoholism show that alcoholism is common among families of alcoholics and that a given alcoholic is likely to have a family history for alcoholism. Moreover, family history, rather than environmental factors, is associated with the development of alcoholism. Contrary to popular belief, histories of deprivation and poverty do not predict alcoholism nearly as well as do biological factors. Furthermore, although mental illness is commonly associated with alcoholism, most alcoholics do not have a mental illness beyond their alcoholism.

Adoption studies reveal that the biological parents, and not the adoptive parents, predict alcoholism in the offspring. Thus, it is the relationship of sons and daughters to parents through blood linkage, and not who raised them or their environment, that predicts the onset of alcoholism. In addition, identical twins who share the same genes (DNA) have a greater prevalence of alcoholism than do fraternal twins who share half of the same genes (DNA), again implicating the biological origin of alcoholism. Importantly, although the prevalence of alcohol dependence typically has been greater in males, studies have shown that the biological risk for developing alcohol dependence is the same among women as among men.

When children are asked for reasons for initiating alcohol (and drug) use, peer influences and exposure to alcohol are the best predictors. Those children who associate with other children who consume alcohol are more likely to try alcohol and use it more often. Other factors, such as poverty and social deprivation, do not predict alcohol use; in fact, certain racial populations such as black males are negatively correlated with onset of alcohol dependence (i.e., they have lower rates of alcoholism). Thus, we can predict alcohol use and onset of dependence with a family history and exposure to alcohol, particularly if an individual is biologically related and exposed to alcohol by peers. Physicians can do well in assessing the risk for onset of alcohol dependence, and who might actually be alcoholic, by inquiring about patients' family histories of alcoholism and patients' association with other drinkers.

2. OTHER DRUG USE ASSOCIATED WITH ALCOHOL DEPENDENCE

Other drug use is commonly associated with alcohol use and dependence. For instance, more than half of those identified as alcoholic will have at least one other drug dependence. This is somewhat age dependent; younger alcoholics are more likely to have another drug dependence. At least 80% of alcoholics are dependent on nicotine, with the vast majority smoking cigarettes. Conversely, 30% of those with nicotine dependence are dependent on alcohol. Cannabis (marijuana) dependence is probably the next most common drug dependence among alcoholics, particularly among younger alcoholics. Most cocaine addicts use alcohol, and many qualify for alcohol dependence, as do most heroin addicts.

Longitudinal studies tend to demonstrate alcohol as a gateway drug, followed by nicotine, for those who go on to develop other drug addictions such as cocaine and heroin. Surveys of high school-age students show that 80 to 90% of students have consumed alcohol during the past year and that 5 to 10% drink daily. Also, the mean age of onset of alcoholism in the U.S. population is 21 years in males and 24 years in females, indicating that alcoholism is a youthful disorder that begins to be prevalent during the teenage years. Initiation of other drug use also occurs during these vulnerable developmental years, when biological and especially peer factors are operative.

3. DEPENDENCE SYNDROME AND SCREENING INSTRUMENTS

3.1. Syndromes

The dependence syndrome, as described in the fourth edition (text revised) of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR), defines a maladaptive pattern of alcohol use leading to clinically significant impairment or distress, as manifested by three or more of the following occurring at any time during the same 12-month period:

- Tolerance (need for increase in dose or loss of effect)
- Withdrawal (typical of alcohol withdrawal: anxiety, dysphonic mood, muscle aches, insomnia, nausea, vomiting)
- Use of larger amounts for longer periods of time than intended
- Persistent desire or unsuccessful attempts to cut down or control alcohol use
- Preoccupation with acquiring alcohol (e.g., frequently in the presence of alcohol; giving up or reducing important social, occupational, or recreational activities due to alcohol use)
- Alcohol use despite knowledge of adverse physical or psychological problems caused or exacerbated by alcohol (e.g., depression; anxiety; restricted options in daily living; pursuit of drugs to the exclusion of social, occupational, and normal interpersonal relationships; poor motivation to solve pain problems through evaluation and treatment; resistance to alternative resolutions to life problems)

Pervasive to these criteria is a loss of control over the use of alcohol, whether an inability to resist initiating alcohol use or an inability to modulate the amount

used during a given episode of alcohol consumption. Accordingly, the dependence syndrome mirrors the following basic definition of behavioral addiction:

- Preoccupation with acquiring alcohol
- Compulsive use of alcohol (continued use despite adverse consequences)
- Pattern of relapse to alcohol use despite adverse consequence

The neuropharmacological basis for behavioral addiction or dependence on alcohol lies in the phylogenetically ancient portion of the brain in the limbic system, which is responsible for instinctual drive states and emotions. The brain center responsible for behavioral addictive use (and dependence) lies in the mesolimbic pathway, which extends from the ventral tegmentum in the midbrain to the nucleus accumbens in the limbic forebrain. Scientific studies show that stimulation of the mesolimbic pathway with alcohol and other drugs reinforces compulsive use of alcohol that closely resembles the out-of-control addictive use seen in patients.

3.2. The CAGE and MAST

As screening instruments, the CAGE (Cutting down, Annoyance by criticism, Guilty feeling, and Eye-openers) and MAST (Michigan Alcohol Screening Test) are useful in identifying problematic use of alcohol and probable alcohol dependence in individuals. In the CAGE, individuals are asked whether they find it hard to cut down, whether others are annoyed by their alcohol use, whether they feel guilty about drinking, and whether they use alcohol as an eye-opener in the morning or early in the day to relieve distress (usually symptoms of withdrawal from alcohol). A positive answer to one question means a possible diagnosis of alcohol dependence, and a positive answer to two questions means a probable diagnosis. However, the CAGE is a screening instrument that requires a follow-up clinical assessment for a revised DSM (DSM-IV-R) or similar objective diagnostic scheme for definitive diagnosis.

The MAST is a self-administered questionnaire that comes in various lengths (e.g., brief, standard), depending on the number of questions indicating a diagnosable alcohol dependence according to the number of positive answers. As with the CAGE, the MAST requires a follow-up clinical interview to confirm the diagnosis of alcohol dependence.

4. ALCOHOLISM TREATMENT: PHARMACOLOGICAL APPROACHES

We know well that pharmacological treatment of withdrawal from alcohol or medical detoxification reduces morbidity and mortality from chronic alcohol use. The mortality from untreated delirium tremens ranges between 20 and 50%, and untreated alcohol withdrawal can complicate other conditions associated with alcohol withdrawal such as cardiovascular problems (e.g., abnormal heart rhythm, hypertension).

The reliable and commonly used medications to treat alcohol withdrawal are from the benzodiazepine class and include diazepam (Valium) and lorazepam (Ativan). Benzodiazepines share cross-tolerance and dependence with alcohol and are effective in suppressing signs and symptoms of withdrawal from alcohol in the brain at the gamma amino butyric acid (GABA) receptors on nerve cells. These are particularly effective medications to treat withdrawal when used in the short term under clinical supervision for alcohol withdrawal.

Moreover, pharmacological management of alcohol withdrawal improves compliance with behavioral approaches to alcohol dependence in promoting abstinence from alcohol. Alleviation of the noxious symptoms of withdrawal, such as anxiety, depression, malaise, and intense cravings for more alcohol, will reduce the risk for continuous alcohol consumption and provide a transition to the abstinent state. Behavioral management is typically not effective in the setting of active alcohol use due to the poor insight and judgment and the reduced motivation in the intoxicated and withdrawal states from alcohol use.

5. PSYCHIATRIC AND MEDICAL COMPLICATIONS

Alcohol regularly induces or causes a plethora of psychiatric symptoms in varying degrees of severity. Depression and anxiety can be expected to occur in virtually any chronic drinker (e.g., daily or weekly consumption). Correspondingly, these symptoms usually subside within a few days following cessation and during abstinence from alcohol. Predictably, chronic drinkers experience sadness, loss of energy, anhedonia, hopelessness, and helplessness that reaches degrees equivalent to those of major depression. In addition, suicidal thinking is common among drinkers; alcoholism is the most common diagnosis among those attempting or completing

suicide. Moreover, 25% of alcoholics commit suicide. Often, these suicidal thoughts subside within days of abstinence from alcohol.

Anxiety is a prominent state during alcohol withdrawal, and most chronic drinkers spend some portion of the day in withdrawal as the blood alcohol level drops. Also, tolerance to alcohol diminishes with chronic intake, such that greater symptoms of anxiety and depression reflect increasing dependence on alcohol. Usually, anxiety and depression subside with abstinence within a few days.

Occasionally, visual and auditory hallucinations, as well as delusions (mostly paranoid in type), can occur in chronic drinkers. Typically, these subside with abstinence, although short-term treatment with anti-psychotic medications may be required.

Importantly, 50% of those with significant psychiatric diagnoses will use alcohol and often qualify for an alcohol dependence diagnosis. Although a presumptive dual diagnosis can be made on the basis of history and mental status examination, abstinence from alcohol over time is often required to make a definitive diagnosis to exclude the alcohol-induced effects on mood and thought.

6. ALCOHOLISM TREATMENT: BEHAVIORAL APPROACHES

Although there are many studies on a variety of behavioral approaches for alcohol problems, these methods are not designed for "typical" alcoholics as defined in the DSM-IV-R. Consequently, brief interventions do not work well in individuals who have actual alcohol dependence with the concomitant loss of control over alcohol use. Moreover, studies of behavioral management approaches typically do not provide long-term follow-up of their effectiveness beyond a few months. These approaches have a variety of names, such as relapse prevention and motivational enhancement, but most are derived from cognitive behavioral techniques. However, one approach termed "12-step facilitation" shows long-term positive results principally due to referral to 12-step groups such as Alcoholics Anonymous (AA).

An approach called "controlled drinking" is a misnomer and has caused substantial confusion, and ultimately harm, to those who suffer from alcohol dependence. In studies on controlled drinking, abstinence was the best predictor of those who experienced the fewest adverse consequences from alcohol use

following controlled drinking exercises. Moreover, the mortality from these controlled drinking techniques during follow-up was high among the study participants. Carefully constructed studies of actual alcoholics revealed that once loss of control over alcohol use manifested itself, the return to controlled drinking was not possible. Thus, abstinence is required to prevent adverse consequences from alcohol once alcohol dependence is established in individuals.

Importantly, behavioral approaches that employ the 12-step approach to the treatment of alcohol dependence show significant abstinence rates from alcohol. These approaches incorporate cognitive behavioral management with a focus on alcoholism as an independent disease that responds to treatment and recovery methods over time. Typically, treatment of alcoholism occurs in inpatient and outpatient settings following detoxification from alcohol and concentrates on increasing the individual's ability to refrain from relapse to alcohol and to maintain the abstinent state.

Outcome studies show that following a treatment experience in a structured inpatient or outpatient treatment program, 60% of patients achieve abstinence at 1 year no matter what follow-up care they may receive. However, they achieve abstinence rates of 70% if they attend continuing care in the treatment program (20–30%) or attend 12-step groups regularly (50% attended at least one AA meeting per week). In addition, they improve their abstinence rates to 80% if they both receive continuing care and attend AA meetings.

7. LONG-TERM RECOVERY FROM ALCOHOLISM

Although spontaneous recovery from alcoholism occurs, the only assurance of long-term abstinence supported by empirical data and clinical experience is continuous attendance in AA. Surveys of its members reveal an 83% probability of abstinence for the next year if an individual is abstinent between 1 and 5 years with regular attendance of at least one AA meeting per week and a 91% probability of an additional year if the individual is abstinent in AA for more than 5 years.

Clinical experience shows that relapse to alcohol use is very uncommon for someone who attends AA meetings regularly, particularly after being abstinent in AA for more than a year. However, for those who relapse, their attendance at AA meetings had ceased for a substantial period prior to the relapse, typically years

before the actual relapse to alcohol use. Also, use of other addicting drugs, such as cocaine, cannabis, heroin, and nicotine, and/or use of addicting medications, such as opiate analgesics and benzodiazepines, render continuous abstinence difficult and reliably predict relapse to alcohol.

8. BENEFICIAL VERSUS HARMFUL EFFECTS OF ALCOHOL

Some studies have indicated a modest but positive effect of reducing the risk of cardiovascular disease or heart attacks in those who regularly consume one to two drinks of alcohol per day. These studies are featured in influential medical journals, implicating alcohol as a therapeutic agent. Despite the overwhelming evidence of the harmful effects of alcohol on health, these studies advocate the use of alcohol without instructing individuals about its untoward or adverse effects.

For instance, studies show that the same amount of alcohol that is described to be beneficial (one or two drinks per day) will significantly reduce intelligence quotients (IQs), resulting in poor attention, poor concentration, and impaired short-term memory. Studies of alcoholics show that chronic alcohol use causes reduced cognitive capacities that tend to reverse with abstinence from alcohol. Importantly, the reduced IQs correlated with the age of the drinker; thus, the intended audience for putative efficacy on incidence of heart disease is those who are at high risk for the dementing properties of alcohol—namely older adults.

9. PREVENTION OF ALCOHOL USE AND PROBLEMS

Currently, our public health policy and controls for alcohol consumption are archaic and counterproductive. The example of advocating alcohol to reduce heart disease in high-risk populations is indicative. Thus, alcohol abuse is not commonly thought of as a health issue but rather is regarded as a legal and moral issue. The policies usually do not support educating young people and adults on the responsible use and the potential pitfalls and risks of alcohol use, particularly by those at high risk. Most discussions focus on the legal age for drinking and the legal limit of intoxication. Not well discussed are that 80% of homicides occur in the setting of alcohol and that alcohol use and dependence is the leading risk factor for suicide at any age.

Despite these well-established adverse consequences and firm knowledge that onset of alcohol use and alcohol problems occur at young ages, there is blatant advertising clearly and convincingly targeted to these high-risk populations. Our controls over exposure to alcohol for vulnerable populations, and our education in elementary, middle, and high schools, are woefully lacking if we are to protect these vulnerable and at-risk populations. Currently, we do not hold producers of alcohol accountable for the high rate of alcohol-related consequences (in the billions of dollars) due to our lack of policy to protect the public against adverse consequences from alcohol use.

See Also the Following Articles

Diagnostic and Statistical Manual of Mental Disorders ▪ Drug Abuse ▪ Drug Dependence

Further Reading

Grant, B. F. (1997). Prevalence and correlates of alcohol use and DSM-IV alcohol dependence in the United States: Results of the National Longitudinal Alcohol Epidemiologic Survey. *Journal of Studies in Alcohol*, *58*, 464–473.

McLellan, A. T., Lewis, D. C., O'Brien, C. P., & Kleber, H. D. (2000). Drug dependence, a chronic medical illness: Implications for treatment, insurance, and outcomes evaluation. *Journal of the American Medical Association*, *284*, 1689–1695.

Miller, N. S., & Goldsmith, R. J. (2001). Craving for alcohol and drugs in animals and humans: Biology and behavior. *Journal of Addictive Diseases*, *20*, 87–104.

Amnesia

Ginette Lafleche and Mieke Verfaellie

Memory Disorders Research Center, Boston University School of Medicine, Boston, Massachusetts, USA

GLOSSARY

anterograde amnesia The inability to learn and recall events and facts following the onset of amnesia.

declarative memory Memory for concepts, facts, and personal events; declarative memories are accessible to consciousness and can be retrieved at will.

episodic memory A component of declarative memory; it refers to memories of personally experienced events, placed in specific spatial and temporal contexts.

nondeclarative memory A collection of unconscious forms of memory, including procedural memory, simple forms of conditioning, and priming.

priming A component of nondeclarative memory referring to the facilitation in processing information as a consequence of prior exposure to that information or to related information.

procedural memory A component of nondeclarative memory referring to the acquisition of skills and habits through repeated practice.

retrograde amnesia The inability to retrieve information about facts and events known prior to the onset of amnesia.

semantic memory A component of declarative memory referring to one's fund of knowledge of facts, concepts, and word meanings; these memories are impersonal and contain no reference to the time and place in which they were formed.

Through the study of amnesic patients, cognitive neuroscience has advanced our understanding of the processes that underlie learning and remembering. Memory is now understood to be a collection of functional systems, with each contributing uniquely to the storage and retrieval of information. The study of neurological populations with lesions that interfere with the function of specific components of memory has revealed distinct roles for these memory systems. This understanding has provided the theoretical foundation for evidence-based treatment approaches.

1. INTRODUCTION

In 1606, Shakespeare's Lady Macbeth described memory as "the warder of the brain," as a function that somehow stands apart from the brain's other cognitions while bearing witness to them. She made this reference while planning to induce a temporary amnesia in the king's attendants by getting them drunk. Her aim in doing so was to ensure that the attendants would not recall the events surrounding her husband's intended murder of the king. She knew that after the

deed was done and their alcohol-induced stupor passed off, they would be left witness to a blank, to a discontinuity in their experience, and not to the murder. For memory to be capable of keeping an account of the swift interplay of sensations, perceptions, and thoughts that make up events and thereby give experience its continuity, memory must itself have an underlying continuity. If it did not, there would be nothing for people to "go back to" that would allow them to witness an earlier event again, and Lady Macbeth would have had nothing to fear from any witnesses to her husband's intended crime. Now, four centuries later, modern cognitive neuroscience is going beyond description. It investigates how memory enables people to actively bring the past into the present moment, thereby allowing them to experience their personal persistence through time.

The fact that memory can fail has been known for millennia, but its failure was not specifically called "amnesia" (from the Greek for forgetfulness) until relatively recently in 1786. In current common use, amnesia refers loosely and interchangeably to either one of two quite different states or to both of them. First, it refers to a state in which memories that were available for recall in the past have been lost and are no longer available in the present. Second, it refers to a state in which current experiences are not being kept account of and duly recorded, with the result that they will not be available for recall in the future. The first usage is reflected in the Oxford English dictionary definition of amnesia as "loss of memory" and roughly coincides with retrograde amnesia, whereas the second usage is reflected in the Merriam–Webster dictionary definition as "forgetfulness" and roughly coincides with anterograde amnesia. The loss of previously consolidated memories, on the one hand, and the inability to create and retain new memories, on the other, are states that can occur separately. However, more often than not, both occur together in the same individual following a single event, hence the tendency to refer to either or both of them without distinction by means of the same term "amnesia."

In addition to its common meaning, the term "amnesia" has both a general sense and a specific sense in current scientific literature. In general, it refers to any pathological loss of the ability to acquire or recall information. In that sense, it encompasses psychogenic amnesia, posttraumatic amnesia following closed head injury, memory loss that occurs along with other major cognitive impairments in progressive degenerative neurological disorders such as Alzheimer's disease,

and isolated global memory loss that characterizes the amnesic syndromes. In its more restricted use, it refers specifically to these latter syndromes, and it is in this sense that the term is used in what follows.

Early investigation of the amnesic syndromes proceeded both by means of clinical case descriptions and through the correlation of lesion analyses with deficits found through neuropsychological investigations. The most widely known of the early case studies was of the patient "H.M.," and it is through studies of his memory function that the current understanding of amnesia began to unfold.

2. THE CONTRIBUTION OF PATIENT H.M.

At 9 years of age, H.M. suffered a head injury followed by seizures that became more frequent and more severe over the subsequent 18 years. In 1953, as a treatment of last resort, he underwent an experimental neurosurgical operation in which a large portion of the medial temporal region of his brain was removed bilaterally. That extensive resection was successful in reducing his seizures, but it left him with an unexpected devastating memory loss.

One striking finding that came to light in the early studies of H.M.'s memory loss was that his profound forgetfulness and loss of personal memories had occurred without any damage to his general intellectual ability, attention, and language. This discovery established that normal medial temporal lobe functioning is necessary for the accumulation of memories. However, the fact that its dysfunction does not interfere with the attention, thoughts, and perceptions that give rise to experiences in the first place suggests that memory is specifically linked to the function of the medial temporal lobe. Further studies of H.M.'s memory impairment led to the current understanding of the characteristics of hippocampal amnesia. They also revealed that memory, rather than being a unitary capacity, is composed of several functional systems, not all of which are impaired in amnesia. Those systems fall into two major classes. The first group involves nondeclarative memory, that is, the form of memory that functions without awareness. The second group involves declarative memory, that is, the form of memory that supports conscious retrieval of experiences. It is the latter form of memory that is impaired in amnesia.

3. IMPAIRED MEMORY IN AMNESIA: DECLARATIVE MEMORY

After H.M.'s operation, it became apparent that he had suffered damage to the memory system that supports the conscious (intentional) retrieval of experiences. Not only did he fail to remember the events of his then current day-to-day life, but he also could not recall memories he had accumulated over much of his life prior to the onset of his amnesia. This clinical profile of severe anterograde amnesia together with an extensive retrograde memory impairment is typical of hippocampal amnesia.

3.1. Retrograde Amnesia

H.M.'s inability to recall memories that he had formed prior to the onset of his amnesia extended back approximately 11 years. The density of his retrograde amnesia followed Ribot's law, a semiquantitative description offered in 1881 that states that the "loss of memory is . . . inversely [related to] the time that has elapsed between any given incident and the fall (injury). . . . The new dies before the old." H.M. showed just such a temporal gradient. His inability to recall autobiographical and public events from shortly before his operation was total, but his amnesia was found to be less dense as testing probed further back in time and he was able to remember more and more.

Investigation of patients with a variety of amnesic syndromes has demonstrated that such a temporally graded retrograde amnesia is common and that its severity and duration tend to vary with the size of the brain lesion. Retrograde amnesia can affect both the recall of autobiographical and public events (episodic memory) and the retrieval of facts and concepts (semantic memory) that were known before the onset of amnesia. However, episodic and semantic memory are differentially affected across syndromes. For example, some patients have severe retrograde amnesia for events together with relatively preserved memory for facts and concepts, whereas others show the opposite pattern with severely impaired semantic memory but relatively preserved episodic remote memory.

The finding of a graded decrease in the density of memory loss as probing moves backward through time starting from the onset of amnesia establishes two important facts regarding how memories are stored and retrieved. First, it establishes that a functional medial temporal region is required for permanent learning of new information. Second, it indicates that the medial temporal lobe is not the ultimate repository of long-term memories given that long-established memories can survive and be available for recall despite the presence of medial temporal damage. It has since been found that long-term memories are stored in the neocortex, specifically in the lateral temporal regions. This is confirmed by the fact that in patients with lesions in these areas, even old memories can be permanently lost, so that their retrograde amnesia is not temporally graded. Thus, the medial temporal lobes and neocortical areas make unique contributions to information storage. When an event is originally experienced, separate aspects of the event, such as its perceptual characteristics, emotional connotations, and associated thoughts, are processed in initially unconnected sites throughout the cortex, as determined by the functional specialization of various parts of the brain. The medial temporal lobes, and especially the hippocampus, provide the glue that allows these diverse aspects to be linked together into a single representation. When a part of any event is recalled again, the hippocampus is engaged, and it simultaneously reactivates the spatially separated cortical traces that together make up the brain's representation of the whole event. With repeated activation over time, links are established among the involved neocortical sites. Eventually, these separate neocortical sites that collectively represent a single event become interlinked to the extent that hippocampal activation is no longer necessary for recall of the event as a whole. It is at that point that the consolidated memory can survive medial temporal lobe damage.

3.2. Anterograde Amnesia

The nature of the impact of H.M.'s neurosurgery on his cognition gradually became apparent when it was noticed that H.M. consistently failed to recall having eaten a meal and that he failed to develop any ability to get to know the people who provided his daily care. This anterograde amnesia was so pervasive that he could hold onto new information for only a few seconds and then only if there was no distraction. If his attention was diverted, the information faded away and was not available for recall. As was noted earlier, the normal functioning of the hippocampus and surrounding medial temporal lobe structures is crucial for forming and storing new memories. In general, the severity of anterograde amnesia is proportional to the extent of

damage to medial temporal structures. Furthermore, because the hippocampus receives input from a variety of neocortical regions that process information in the various sensory modalities, such anterograde loss is invariably global; that is, it involves both verbal and nonverbal information in all possible modalities of presentation. Such dense global anterograde memory impairment is the hallmark of medial temporal lobe amnesia.

Anterograde amnesia is evaluated in the clinic through tasks that measure recall and recognition of information to which the individual is exposed during a testing session. Recall tasks require that recently learned information be consciously and deliberately retrieved in response to a question, whereas recognition tasks require that previously encountered information be distinguished from, and so recognized among, foils. The recognition format is generally easier for normal individuals because it requires less retrieval effort than does recall. Recognition is thought to be easier for individuals with intact memories because recall depends on conscious and effortful recollection, whereas recognition can be supported not only by recollection but also by familiarity. Familiarity refers to the sense of knowing that arises when a stimulus is processed with ease. Recollection and familiarity can be distinguished from one another experientially. For example, consider the situation where a man hears his name being called out while he is in a crowd. He turns and sees the woman who called his name and has a short conversation with her. The woman is familiar to the man, but he cannot recall any context that accounts for that familiarity. Several days later, the man receives a telephone call from the woman in which she identifies her place of work and, at that moment, he knows who she is and is flooded with memories of details and events. Sudden recollection replaces felt familiarity. In the laboratory, patients with amnesia secondary to medial temporal lobe damage generally show impairments on both recall and recognition tasks, suggesting that familiarity and recollection both are impaired. However, some amnesic patients show relatively preserved recognition memory, and this may reflect their residual ability to use familiarity as a basis for remembering.

Patients with lesions involving the frontal lobes, either directly or indirectly, can also have severe deficits in memory. However, although their performance is impaired on tasks of free recall, they show normal performance on tasks that require the recognition of previously learned information. This performance disparity between recall and recognition is due to the distinct roles of the medial temporal lobes and the frontal lobes. The medial temporal lobes contribute to the encoding, storage, and retention of newly learned information, whereas the frontal lobes support the organizational and strategic aspects of memory necessary for developing encoding and retrieval strategies, monitoring and verifying memory output, and setting order within the recalled memories.

4. PRESERVED ASPECTS OF MEMORY IN AMNESIA: NONDECLARATIVE MEMORY

H.M.'s severe amnesia notwithstanding, some aspects of his memory functioning were found to be intact following his surgery. In contrast to his inability to consciously retrieve information acquired before and after onset of his amnesia, it was found that H.M. showed evidence of learning on tasks that did not require his awareness of the learning episode.

4.1. Procedural Memory

Careful experimental study brought to light that H.M. was able to acquire specific perceptual–motor skills as well as do normal individuals. One such motor task involved his tracing the outline of a star with a stylus while gaining feedback of his ongoing performance through the reflection of his performance in a mirror. H.M. improved with repetition exactly as did normal individuals. But for him, each trial was experienced as if it were his first trial, despite his steady improvement. He had no way of knowing that he was getting better at it, as normal people did. This finding was one of the first to show that procedural memory (i.e., the ability to know how) is distinct from declarative memory (i.e., the ability to know that something occurred). It also established that procedural memory does not depend on the integrity of the medial temporal lobes. Procedural memory develops gradually through repetition without any requisite awareness that the skill is being learned.

4.2. Repetition Priming

Priming is another form of learning that occurs without conscious awareness. It is defined as a bias or facilitation in identifying or responding to certain information to which one has been exposed previously. In contrast

to procedural learning, the effect of priming can be seen following a single exposure. In a typical priming task, an individual is asked to view a list of words or pictures. The individual is then asked to perform a seemingly unrelated task such as naming words or identifying degraded pictures. Unbeknownst to the individual, some of the stimuli in the second task were part of the list of stimuli that were seen in the first task, whereas others are presented for the first time. Priming is in evidence when naming or identification of previously seen stimuli occurs faster or more accurately than does identification of stimuli not seen previously. Amnesic patients show intact priming across a broad variety of tasks. Sometimes priming reflects facilitation of the processes that support the perception of stimuli, whereas other times priming reflects the facilitation of processes that support conceptual analysis (i.e., meaningful interpretation) of stimuli. Both perceptual priming and conceptual priming are independent of medial temporal structures, and both depend on distinct neocortical regions.

5. AMNESIC SYNDROMES

Amnesia can result from a variety of etiologies that cause damage to various regions of the brain. For example, diencephalic amnesia, the memory disorder of patients with Korsakoff's syndrome, arises from damage to structures that make up the diencephalon, particularly the mammillary bodies, the anterior thalamic nuclei, and the medial dorsal thalamic nucleus. Basal forebrain amnesia, in contrast, arises in patients with a history of rupture and repair of an anterior communicating artery (ACoA) aneurysm and is caused by damage to structures of the basal forebrain, the septum, and frontal brain regions. The following is an overview of the clinical characteristics of five of the more common amnesic syndromes.

5.1. Herpes Simplex Encephalitis

Herpes simplex encephalitis (HSE) is an acute inflammation of the brain caused by a herpes simplex virus infection. Some patients recover fully following infection, but most are left with a cluster of cognitive deficits that include a memory disorder. The heterogeneous pattern of these cognitive impairments reflects the extent and variability of damage to the brain. Among those patients who recover, a small number are left with a circumscribed amnesic syndrome similar to that of H.M. The fact that these two forms of amnesia share a clinical

profile is not surprising given that the herpes simplex virus preferentially affects medial temporal brain structures, including the hippocampus and adjacent entorhinal, perirhinal, and parahippocampal cortices. In all postencephalitic patients, the extent of any anterograde memory impairment is proportional to the amount of damage to the medial temporal lobes. Although some patients are incapable of any new learning, others can learn and can benefit from increased study time, external cues, and/or repeated exposure. Because the herpes virus can affect the brain asymmetrically, the pattern of anterograde memory loss will also depend on the laterality of damage. The greater the damage to the right temporal region, the greater the difficulty in performing nonverbal/visual memory tasks (e.g., memory for faces), whereas the greater the damage to the left temporal region, the greater the deficit in verbal memory.

Postencephalitic patients with damage extending to the lateral temporal lobes, the region where memories are permanently stored, will have dense retrograde amnesia for autobiographical and public information with little or no temporal gradient. Damage to the right anterior temporal region interferes primarily with retrieval of autobiographical memories, whereas damage to the left temporal cortex impairs semantic memory.

5.2. Anoxia

Amnesia can result from damage to, or death of, brain tissue due to a lack of oxygen supply to the brain. This can be caused either by reduced blood flow, such as in cardiac arrest or strangulation, or by normal perfusion with hypoxic blood secondary to respiratory distress or carbon monoxide poisoning. When the oxygen supply to brain tissue is disrupted, compensatory mechanisms that maintain cerebral homeostasis are triggered immediately. These protective autoregulatory mechanisms, although effective in adjusting for sudden short-lived changes, will eventually fail in the face of sustained oxygen deprivation. Oxygen deprivation for 3 to 8 minutes will trigger the release of excitatory neurotransmitters that can result in damage to the hippocampus. Shorter events may also damage regions that are perfused by terminal vascular branches or that lie in the watershed regions because these areas are deprived of oxygen early on. Longer lasting events will result in neuronal damage that extends to the cerebellum, the basal ganglia, the thalamus, and neocortical areas. The outcome following an anoxic event depends on many factors, including the cause and duration of the event as well as the age and health status of the individual.

Therefore, it is not surprising to find variability in the clinical profiles associated with anoxic events.

Most individuals who have suffered an anoxic brain injury experience a memory disorder along with other cognitive impairments. On occasion, an isolated amnesic syndrome is documented secondary to a lesion in the hippocampus. When the lesion is limited to the CA1 area of the hippocampus, there is moderately severe anterograde amnesia together with very mild retrograde loss. When the lesion extends beyond the CA1 area of the hippocampus but remains limited to the hippocampal formation, there is severe anterograde memory impairment and more robust retrograde memory impairment that can extend back more than 15 years. More commonly, however, anoxic patients have a memory impairment that is akin to that seen in patients with damage to the frontal brain region due to disruption of frontal–subcortical circuits as a result of damage to watershed zones in the cerebral cortex. Rather than encoding, storage, and retention deficits, these patients demonstrate impairments at the level of the organizational and strategic aspects of memory.

5.2.1. Anterior Communicating Artery Aneurysm

Memory deficits, along with associated behavioral disorders, frequently follow rupture and surgical repair of an ACoA aneurysm. Such deficits vary from a very mild impairment to a severe amnesic disorder called basal forebrain amnesia. This wide variability reflects the fact that the ACoA perfuses a broad anatomical brain region, all of which is vulnerable in the event of rupture to damage from infarction, either directly or secondary to subarachnoid hemorrhage, vasospasm, or hematoma formation. Basal forebrain amnesia is usually due to damage to the septal nucleus and the subcallosal area. Disruption of hippocampal functioning may play a role in the amnesia of at least some patients with ACoA aneurysm because several basal forebrain nuclei contain a large number of cholinergic neurons that innervate the hippocampus as well as large neocortical regions. Disruption of frontal network systems can occur and may also contribute to the quality of the memory deficit seen in some ACoA aneurysm-related forms of amnesia.

The severe deficit in recall found in basal forebrain amnesia shares superficial similarities with the anterograde memory impairment that characterizes medial temporal lobe amnesia. However, the deficit in basal forebrain amnesia is due to inefficient encoding and not to the failure of consolidation that characterizes medial temporal amnesia. As a result, ACoA patients will benefit from the use of encoding strategies, whereas patients with medial temporal lobe damage will not. Another difference is found in the performance of ACoA patients on recognition tasks, where they are found to succeed more often than their medial temporal lobe counterparts. This recall–recognition disparity is attributed to the frontal executive component of the amnesia that follows ACoA aneurysm, particularly the disruption of strategic effortful search processes that give access to information stored in memory. Another executive contribution to the disorder is the tendency of some ACoA patients to score many false positives on recognition tasks, a finding that suggests an additional impairment in the ability to monitor the outcome of a memory search. Retrograde memory is nearly always impaired in amnesia secondary to ACoA aneurysm, and it invariably shows a temporal gradient. However, patients do appear to benefit substantially more from cueing than do other amnesic groups, suggesting that impaired retrieval efficiency contributes to both the anterograde and retrograde aspects of their amnesia.

5.2.2. Stroke

Amnesia is a common consequence of infarction of the posterior cerebral artery (PCA). It results from neural tissue damage caused by the interruption of blood flow to a large part of the medial temporal lobes, particularly the posterior two-thirds of the hippocampus, the parahippocampal gyrus, and other critical pathways that connect the hippocampus to surrounding brain areas. A more posterior extension of the lesion will result in other neuropsychological deficits in addition to the memory disturbance, for example, visual field defects and other visual disturbances that may affect reading and cause problems with color identification, space perception, and/or object naming. The typical memory disturbance associated with bilateral PCA infarction is an inability to establish new memories in the presence of preserved intelligence and attention. Retrograde memory problems are also often present. There have been several reported cases of bilateral PCA infarction that spared the medial temporal lobes proper but that involved the occipital lobes bilaterally as well as the deep white matter of both the occipital and temporal lobes. These patients present with a visual amnesic syndrome that results from the disconnection between occipital cortices involved in visual processing and temporal brain regions supporting memory. There have been other reported cases in which the PCA

infarction was unilateral. Patients with infarction of the left PCA present with a selective verbal memory deficit, whereas patients with infarction of the right PCA have preserved verbal memory but impaired visual processing skills and impaired visual memory.

Amnesia can also be caused by thalamic infarction. In such cases, the severity of the memory impairment is related to the site of damage within the thalamus. Lesions that damage the mammillo–thalamic tract, in particular, have been associated with severe anterograde amnesia. Infarction of the medial dorsal thalamic nuclei has also been associated with memory impairments, but it appears that the damage must extend beyond the medial dorsal nucleus to include the mammillo–thalamic tract or anterior nucleus for the development of a severe amnesic disorder. Because the thalamus has rich connections with the frontal lobes, this anterograde amnesia is also accompanied by an increased sensitivity to interference and by impairments in executive functioning. As with other amnesic syndromes, left-sided lesions result in impairments on tasks of verbal learning, whereas right-sided lesions result in nonverbal/visual memory impairments. Retrograde memory deficits following thalamic infarction are variable; some patients are found to have little impairment in remote memory, whereas others demonstrate severe long-term memory impairments.

5.2.3. Wernicke–Korsakoff Syndrome

Wernicke–Korsakoff syndrome (WKS) is seen in patients with a history of long-term alcohol abuse in association with poor nutrition and a lack of Vitamin B1 (thiamine). In acute Wernicke's encephalopathy, patients exhibit confusion, a gait disorder (ataxia), and abnormal eye movements (oculomotor palsy). Treatment with large doses of thiamine may result in improvement in, or even reversal of, some of these symptoms. However, most patients are left with a permanent dense amnesic disorder referred to as Korsakoff's syndrome. This amnesic syndrome arises from damage to the thalamic nuclei, the mammillary bodies, and the frontal system.

Patients with WKS suffer from both anterograde and retrograde amnesia. Several explanations have been proposed to account for their episodic memory impairment. Although early models emphasized their superficial and deficient encoding strategies or their failure to inhibit competition from irrelevant material at the time of retrieval, current views agree that an explanation of their learning deficits is best accounted for by a theory that integrates both encoding and retrieval processes.

The retrograde amnesia in WKS has a steeper temporal gradient than that found in medial temporal lobe amnesia. The concomitant presence of frontal dysfunction in Korsakoff's patients is believed to account for their poorer performance on remote memory tests. Intelligence is usually preserved in Korsakoff's patients, but there are often associated cognitive and neurobehavioral deficits that are unique to this patient population. In particular, some combination of impaired planning and initiation, passivity, apathy, confabulation, and limited insight is nearly always found. These symptoms are thought to arise from associated frontal dysfunction.

6. NEUROPATHOLOGY OF AMNESIA

For centuries, any lapse in the functioning of the warder of the brain's cognitions was described, and eventually defined, as simple forgetfulness. During the first half of the 20th century, it was thought that memory, as the cognition that keeps account of other cognitions, was not dependent on the activity of a circumscribed brain region (as are those other cognitions) but rather was directly dependent on the functioning of the whole brain. There also emerged an opposing view that held that memory function was localized in the brain. There was no conclusive experimental evidence for either view until, at midcentury, the consequences of H.M.'s surgical resection decided the matter in favor of a localized brain representation for declarative memory. Later evidence similarly established distinct localizations for priming and procedural learning, two forms of nondeclarative memory. Cognitive neuroscience has since clarified the nature of those localizations in substantial detail.

6.1. Declarative Memory

Declarative memory is mediated by a group of interconnected structures that are part of the extended limbic system. Within the limbic system, two interacting memory circuits can be identified: the Papez circuit and the basolateral circuit. The Papez circuit is composed of the hippocampus, the fornix, mammillary bodies, the anterior thalamus, and the posterior cingulate gyrus (with additional connections to the basal forebrain via the fornix). The basolateral circuit is composed of the amygdala and surrounding perirhinal cortex, the dorsomedial thalamus, and the prefrontal cortex. Although damage to any part of either circuit in isolation will impair memory

function, damage to both circuits will result in a profound amnesic disorder. The contribution that each circuit makes to memory function remains a matter of active debate. On one side of the discussion are those who believe that both circuits are involved in all aspects of declarative memory. Lesions affecting both circuits would, therefore, be expected to result in more severe amnesia simply because more of the relevant neural tissue would be dysfunctional. Others suggest that the two circuits make qualitatively different contributions to memory. In particular, they suggest that the hippocampus and related structures in the Papez circuit support the recollection of episodic information, whereas the perirhinal cortices and related structures in the dorsolateral circuit support judgments of familiarity. In this case, damage to both circuits would impair more functions than would damage to either circuit alone and, therefore, would explain current clinical findings. Future studies of patients who present with selective lesions will be needed to resolve this debate.

The finding of temporally graded retrograde amnesia in association with damage to the medial temporal lobe reveals that this brain region plays a critical role in the establishment of memory and also suggests a subsequent slow transfer of memory to other brain regions. The finding that old memories, both autobiographical and semantic, are left untouched by damage limited to the hippocampus suggests that memories are not stored there. Long-term storage takes place in neural networks in the neocortex.

6.2. Nondeclarative Memory

Nondeclarative memory systems are supported by widely varying brain regions, depending on which sensory mode is involved in a given task and whether or not performance of the task involves higher associative functions. For example, evidence from neuroimaging studies, together with clinical data from individuals who have suffered focal cortical damage, has established that priming finds its substrate in the neocortex. Specifically, the substrate for perceptual priming is the relevant unimodal cortex (i.e., visual priming in the occipital visual cortex, auditory priming in the auditory cortex), whereas the substrate for conceptual priming is located in multimodal association cortices. The same brain regions that are involved in the initial processing of information are also involved in the more fluent processing that follows repetition, and priming is accompanied by a reduction in neural activity in these regions.

Although priming manifests after a single exposure to a stimulus, procedural memory requires a series of repetitions for its manifestation. Consequently, the involvement of relevant brain regions and their changes over time is more complex for procedural memory. The initial acquisition of motor skills engages the motor/prefrontal cortices, basal ganglia, and the cerebellum. Over time, however, two opposite tendencies are evident. On the one hand, as procedural learning develops and a skill can be performed with less effort, the prefrontal cortex and cerebellum gradually become less activated. This repetition suppression effect mirrors that seen in the neural substrate for priming. On the other hand, repetition leading to increased skill engages the higher order motor cortex more than it had initially been engaged. This latter finding indicates expanded cortical involvement in the retention of procedural memory for motor tasks that mirrors the increased neocortical involvement in the long-term retention of declarative memories.

7. APPROACHES TO TREATMENT OF AMNESIC DISORDERS

The treatment of memory disorders aims at enhancing day-to-day memory functioning and routine so as to increase an individual's level of independence. The choice of treatment will depend on both cognitive and noncognitive factors. Noncognitive factors to be considered include psychosocial context (e.g., family situation, educational background, lifestyle habits) and emotional factors (e.g., level of insight, motivation, neuropsychiatric symptoms). Of critical importance among these factors are level of insight and motivation. Research has shown that interventions are unsuccessful in patients who fail to appreciate that their memory is impaired and/or who are unmotivated. In these patients, efforts should initially focus on increasing level of insight and/or motivation. Cognitive factors that must be taken into account include premorbid abilities and skills and postmorbid neuropsychological deficits, including a clear delineation of those areas of memory that are impaired and preserved. Within the context of a holistic individualized approach, a memory remediation program can be developed based on current understanding of the processes that support learning and memory.

Several treatment approaches for densely amnesic patients capitalize on nondeclarative memory

processes because this form of memory remains intact in most patients. The "vanishing cues" technique is an example of a treatment method that recruits preserved implicit perceptual memory processes to teach patients domain-specific facts or concepts. The technique takes advantage of patients' preserved ability to complete studied items in response to word fragment cues. In a typical vanishing cue paradigm, patients are given a definition and are then presented with as many letters as is necessary to produce the target word. With training, the letter cues are gradually reduced until the patients can spontaneously generate the sought after information. Success has been achieved using this technique in teaching patients computer-related vocabulary, business-related terms, and novel concepts. Learning by means of this technique is slow and laborious but can lead to surprisingly good retention, particularly if the information to be learned uses a knowledge base that is already familiar to the individual. However, a caveat arises from the inherent reliance of the vanishing cues method on the perceptual cues given during learning. As a result, generalization has often been limited, and benefits have been found to be best when the information is used in situations similar to those where learning occurred.

Attention to training contexts may be important when using techniques that take advantage of preserved implicit memory processes. Because amnesic patients have no recollection of the learning episodes, they fail to remember their mistakes and consequently fail to benefit from ongoing error correction. Instead, incorrect responses made during learning are often unconsciously repeated, leading to errors becoming primed and more likely being repeated subsequently. To avoid the perpetuation of errors through priming, some investigators have emphasized the importance of "errorless learning" for patients who have explicit memory impairments. In errorless learning, the possibility of making errors is eliminated by using cues and prompts or by providing the correct answer. The approach has met with some success in teaching memory-impaired patients both new skills (e.g., use of a memory book, programming of an electronic organizer) and new knowledge (e.g., learning of new words). Errorless learning is thought to operate by strengthening residual explicit memory, either alone or together with implicit memory. Its applicability as a method to facilitate learning appears to be broad and promising because errorless learning principles can be applied to a variety of remediation methods.

Other treatment approaches capitalize on preserved procedural learning. Through repetition, skills and habits that are important for activities of daily living or occupation can be taught. Such skills can range from simple assembly tasks to more complex multistep tasks such as learning to type. A variety of compensatory aids (e.g., notebooks, scheduling books, diaries, alarms) and augmentative technologies (e.g., computers, personal digital assistants, paging systems) rely on procedural memory. The notebook is an example of a low-tech aid. It is usually created at the early stage of rehabilitation and contains sections aimed at drilling overlearned personal information (e.g., date of birth, age, address), information about immediate family members (including their telephone numbers), daily schedules, and a daily record of activities. The book is tailored to the individual and can be gradually increased in complexity. Electronic organizers, the most favored external memory aid among normal individuals, are now also being used by memory-impaired patients. Training in the use of such technology requires lengthy practice sessions, within and outside the rehabilitation environment, to foster generalization. Because the acquisition of new skills is time-consuming for everyone involved, careful consideration needs to be given as to whether an electronic device is appropriate for an individual before investing the time and effort. Factors that would argue against it include dense amnesia associated with poor insight, lack of initiative, impaired visual attention, poor motor control, and limited problem-solving skills. Patients who are good candidates for this technology are generally younger, have experience in using electronic devices, and have achieved higher educational levels. Devices used premorbidly are preferable because familiarity increases the likelihood that they will be used effectively outside the clinic.

The preceding examples illustrate approaches that focus on preserved memory systems to teach new skills and habits. Another approach focuses on enhancing impaired forms of memory to improve day-to-day episodic memory by means of internal strategies. Internal strategies require awareness of the learning method and recall of the strategy itself; therefore, they are of limited use to patients who are moderately or densely amnesic. However, they are useful for patients who have mild memory deficits secondary to impaired effortful encoding or retrieval, who have good awareness of their deficits, and who have adequate motivation. These patients are more likely to generalize their training to situations that go beyond the clinic setting.

Examples of internal strategies include mental retracing, feature–name association, and verbal elaboration by means of a story or an association—all of which are skills that promote the use of imagery. The choice of technique will depend on the memory process that is targeted for remediation. For example, techniques that focus on strengthening encoding, and therefore storage, are effective at remediating consolidation problems. Story elaboration is effective in linking together a list of unrelated words through the development of a scenario that features the target words. The use of verbal associations is often effective in recalling a surname. In all of these instances, repeated use of the strategy is important and spaced repetitions, at different times and in different contexts, generally increases the likelihood that information will be learned and become conceptually integrated within a matrix of old memories. Strategies that are most effective when the deficit is at the level of strategic processes that enhance encoding are those that increase the organizational structure of incoming information. For example, learning how to "chunk" incoming information is helpful in streamlining and organizing that information. Organizing information according to themes or categories can also structure learning so that a thematic cue can serve to trigger recall when necessary.

These remediation methods are guided by knowledge regarding the cognitive processes that support learning and memory. More than 50 years ago, the early clinical findings with H.M. and other patients informed the theoretical understanding of the functional systems that comprise human memory. Current cognitive neuroscience has evolved from those early findings and now, in turn, can inform clinical approaches to remediation that enable amnesics to function more effectively in their daily lives.

See Also the Following Articles

Aging and Culture ■ Cognitive Aging ■ Cognitive and Behavioral Interventions for Persons with Dementia ■ Dementia in Older Adults ■ Posttraumatic Disorders

Further Reading

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and Brain Sciences, 22,* 425–489.

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences of the United States of America, 91,* 7041–7045.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus?. *Nature Review Neuroscience, 2,* 51–61.

Corkin, S. (1984). Lasting consequences of bilateral medial temporal lobectomy: Clinical course and experimental findings in H.M. *Seminars in Neurology, 4,* 249–259.

Moscovitch, M., Goshen-Gottstein, Y., & Vriezen, E. (1994). Memory without conscious recollection: A tutorial review from a neuropsychological perspective. In C. Umilta, & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious processes in cognition* (pp. 619–660). Cambridge, MA: MIT/Bradford Press.

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., & Amaral, D. G. (1998). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience, 16,* 5233–5255.

Sagar, H. J., Cohen, N. J., Corkin, S., & Growdon, J. H. (1985). Dissociations among processes in remote memory. In D. S. Olton, E. Gamzu, & S. Corkin (Eds.), *Memory dysfunctions: An integration of animal and human research from preclinical and clinical perspectives* (Vol. 4, pp. 533–535). New York: Academy of Sciences.

Scoville, W., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20,* 11–21.

Squire, L. R., & Kandel, E. R. (1999). *Memory: From mind to molecules.* New York: W. H. Freeman.

Tulving, E. (1998). Neurocognitive processes of human memory. In C. von Euler, I. Lundberg, & R. Llinás (Eds.), *Basic mechanisms in cognition and language* (pp. 261–281). Amsterdam: Elsevier.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus, 8,* 198–204.

Verfaellie, M., & Keane, M. M. (2002). Impaired and preserved memory processes in amnesia. In L. R. Squire, & D. L. Schacter (Eds.), *Neuropsychology of memory* (3rd ed.). New York: Guilford.

Anxiety and Optimal Athletic Performance

John S. Raglin

Indiana University, Bloomington, Indiana, USA

GLOSSARY

anxiety An emotion consisting of unpleasant thoughts and sensations as well as physical changes; it is a response to a situation or stimulus perceived to be threatening or dangerous.

arousal Traditionally defined as a generalized and undifferentiated physiological response to a stressor that is closely associated with negative emotions such as anxiety; research indicates, however, that the physiological responses to standardized stressors can differ significantly across individuals and are often unrelated to negative emotions such as anxiety.

Of all the psychological variables implicated in sport performance, anxiety is regarded to have the greatest impact. Putative explanations for why anxiety can degrade sport performance include diminished cognitive resources, a narrowing of the visual field, diminished motivation, and excessive muscle contraction or co-contraction of opposing muscle groups that impairs coordination or results in physical injury. As a consequence, a wide variety of techniques have been used to assist athletes in controlling or reducing anxiety. Yet in spite of the general consensus within the field of sport psychology that anxiety harms sport performance, there are major disputes as to which theory best describes this relationship and as to what the most appropriate means are to measure anxiety in athletes. This article presents an overview of anxiety and sport performance literature. Limitations of traditional theoretical explanations are described, as are more recent sport-specific theories based on the premise of individual differences in the responses of athletes, with particular emphasis on the Individual Zones of Optimal Functioning (IZOF) model.

1. INTRODUCTION

Anxiety has been defined as an emotion consisting of dysphoric thoughts, unpleasant sensations, and physical changes that occur in response to a situation or stimulus perceived to be threatening or dangerous. According to most theories, anxiety consists of state and trait components. State anxiety indicates the intensity of anxiety experienced at a given moment and can fluctuate widely in intensity over a short time span.

Trait anxiety is a more stable factor that assesses the general tendency of an individual to experience elevations in state anxiety when exposed to stressors such as sport competition. Persons scoring high in trait anxiety should experience greater increases in state anxiety when exposed to a stressor than should persons scoring low in trait anxiety. The experience of anxiety as a response to a stressor such as sport competition is contingent on both an individual's perception of the stimulus and his or her ability to effectively cope with it. Because of this, sport competition may be perceived as threatening to some individuals, neutral to others, and enjoyable to still others.

1.1. Assessment of Anxiety

In sport psychology research, a variety of approaches have been used to quantify anxiety, including the observation of overt behavior, biological activity (e.g., galvanic skin activity, heart rate, stress hormones), and self-reports. No single method is entirely reliable. Assessments of behaviors implicated in anxiety (e.g., pacing) may be an anxiety-reducing strategy for some individuals or may be entirely unrelated to anxiety in other instances. Physiological variables that have been used as biological correlates of anxiety (e.g., electromyogram [EMG]) may be difficult to assess prior to competition, or they may provoke an increase in anxiety in some cases (e.g., sampling blood to assess stress hormones).

Because of these problems, anxiety is most commonly determined by means of self-report questionnaires. In sport research, the most frequently used general measure of anxiety has been the State–Trait Anxiety Inventory (STAI), a 40-item questionnaire that assesses both state and trait anxiety. Despite its proven validity, the efficacy of the STAI and other general measures of anxiety in the context of athletics has been questioned, leading to the creation of more than 30 sport-specific anxiety scales. Among these, the most widely used is the Competitive State Anxiety Inventory–2 (CSAI-2), a multidimensional anxiety measure that assesses self-confidence, somatic anxiety, and cognitive anxiety.

Despite advantages such as the ease of administration and interpretation, self-report measures are not without limitations. The validity and reliability of self-reports are delimited by verbal ability and self-awareness of emotional states. Administering questionnaires near the time of competition can be impractical or disruptive and might even result in increased anxiety by directing attention to internal emotional states. A more serious problem is response distortion, which occurs when individuals respond falsely to questionnaires for reasons such as social desirability, the demand characteristics of the experiment, and personal expectations. Response distortion can be detected through the use of lie scales, but this form of control is rarely used in sport psychology research.

2. TRADITIONAL THEORETICAL PERSPECTIVES ON THE ANXIETY– PERFORMANCE RELATIONSHIP

It has been a long-standing belief in sport psychology that high levels of anxiety experienced during competition are harmful for performance and, if unabated, may even result in some athletes quitting their sport. A variety of interventions have been employed by sport psychology practitioners to reduce anxiety, including hypnosis, progressive relaxation, visualization, biofeedback, autogenic training, meditation, negative thought stopping, and confidence enhancement. However, it also has been posited that anxiety can facilitate performance under particular conditions. This perspective originally stemmed from drive theory, otherwise known as Hullian theory. According to Hullian theory, performance is a function of drive (i.e., physiological arousal or anxiety) and habit strength (i.e., skill). High levels of anxiety should increase the likelihood of correct behavior for well-learned skills, as would be the case for an emotional pep talk presented to a group of talented athletes. Evidence for drive theory in sport settings is lacking, however, and the theory currently has little status in the field of sport psychology.

Theoretical explanations in which high anxiety adversely influences performance have a higher standing in sport psychology, none more so than the Yerkes–Dodson law, familiarly known as the inverted-U hypothesis. The hypothesis stems from the classic work by Yerkes and Dodson, who in 1908 examined the influence of stimulus intensity on habit formation in experiments where mice were timed in maze running. Discrete levels of difficulty were created by manipulating the level of illumination of the maze and subjecting the mice to several intensities of stimulation via electrical shocks. The highest intensity shocks were found to slow learning under the most difficult (i.e., dimmest) maze trial, suggesting that moderate stimulation was best for such conditions. These results have since been widely reported in both general psychology and sport psychology textbooks, and they have been generalized to

a number of constructs such as drive, motivation, learning, arousal, and anxiety.

In sport psychology, the hypothesis is presented as a relationship between athletic performance and either arousal or, more commonly, anxiety. Optimal performance should occur when anxiety is within a moderate range of intensity, whereas deviations above or below this range should result in progressively worsened performance. Hence, anxiety and performance exhibit a relationship describing the shape of an inverted U. In basic terms, optimal performance is most likely to occur when anxiety is neither too high nor too low, but because of the stressful nature of sport competition, it is assumed that it is far more likely for athletes to experience too much anxiety.

2.1. Sport Modifications

The inverted-U hypothesis has been adapted to account for sports with different physical requirements (e.g., fine motor skill vs gross motor skill) and the expertise of the athlete performing the task. Sport tasks that require precise motor control and minimal physical effort (e.g., putting in golf) are posited to benefit when anxiety or physiological arousal is low prior to and during the tasks. As tasks require greater physical effort but less fine motor control (e.g., tackling in football), progressively higher anxiety intensities should enhance performance. The second modification predicts that as expertise and talent in sport tasks increase, the optimal range of anxiety will also be higher compared with athletes who are either beginners or intermediate in skill. It is assumed that more talented athletes possess the necessary motor skills and coping strategies to harness higher anxiety, whereas less skilled or experienced individuals should exhibit worsened performance at the same anxiety intensity. Given information about sport and skill level for a given athlete, it should then be possible to establish an inverted-U function for the individual.

Despite the continued popularity of the inverted-U hypothesis, reviews of the literature in general psychology and sport psychology have concluded that its empirical support is weak or even entirely absent. It has been concluded that much of the research supporting the hypothesis has little bearing on sport because it was conducted in laboratory environments rather than realistic sport settings or because it used nonathletes as test participants. Studies of the inverted-U hypothesis have also failed to support the propositions that optimal anxiety is altered by the motor skills required for

a sporting event or that comparably skilled athletes competing in the same sport benefit from similar anxiety levels.

Another concern expressed in reviews of the inverted-U hypothesis literature and other traditional theories is the assumption that arousal and anxiety are closely related or synonymous constructs. Arousal was originally defined as a global physiological response to a stressor that is closely associated with negative emotions such as anxiety. Subsequent research, however, has shown arousal to be a far more complex phenomenon. Standardized stressors evoke physiological responses that often exhibit little or no intercorrelation, and the pattern of physiological activation can vary considerably across individuals and situations. Physiological variables commonly employed as indicators of arousal (e.g., heart rate, respiration rate) are only weakly associated with self-report measures of anxiety, and this also is true for sport-specific measures of perceived arousal or somatic anxiety. Despite these findings, the conceptualization of arousal as an undifferentiated physiological response closely associated with anxiety persists, particularly within the field of sport psychology.

Finally, reviews of both the general and sport literature conclude that the inverted-U hypothesis and other traditional explanations indicate that athletes should respond uniformly to anxiety. They do not allow for the occurrence of interindividual differences to anxiety despite the fact that it has long been recognized that some athletes perform optimally under high intensities. In 1929, for example, the pioneering U.S. sport psychology researcher Coleman Griffith wrote, "Some of the most distressing cases of anxiety and fear in the dressing room have led to outstanding achievements during the game."

3. SPORT-SPECIFIC THEORIES OF ANXIETY AND PERFORMANCE

3.1. The IZOF Model

The lack of efficacy of traditional anxiety–performance theories, as well as the acknowledgment that individual differences contribute to this relationship, has spurred the development of sport-specific theories that incorporate this concept. Among these theoretical explanations, the Individual Zones of Optimal Functioning (IZOF) model is believed to have the strongest empirical basis. The IZOF model was developed by the Russian psychologist Yuri Hanin from studies of athletes in a wide variety of sports and competitive settings where anxiety was

Basic Tenets of the IZOF Model of Anxiety and Athletic Performance

1. Each athlete possesses an anxiety zone that is associated with optimal sport performance.
2. The optimal anxiety zone may exist anywhere on the continuum of anxiety intensity, from low to extremely high.
3. The optimal anxiety zone is not systematically influenced by either the nature of the sporting event or the skill and expertise of the athlete.

assessed using the Russian-language version of the STAI. The results of Hanin's research supported an ideographic explanation, in which anxiety was associated with performance at the level of the individual athlete, rather than a nomothetic explanation, in which an entire team would respond similarly to anxiety. According to the IZOF model, each athlete possesses an optimal anxiety zone or range that is beneficial for performance. This optimal range may exist anywhere within the continuum of anxiety, from as low to as high as is measurable. When anxiety values fall outside the optimal zone, performance should decrease. The IZOF model further posits that the intensity of the optimal anxiety zone is not predictably influenced by either the type of sporting event or the skill of the individual athlete. As a result, an alternative method is needed to identify a more appropriate required anxiety level for an athlete than has been employed in traditional theories where the responses of athletes to anxiety are assumed to be more uniform (Table I).

3.2. Determining Optimal Anxiety

Two techniques have been developed to identify the optimal anxiety zone of an athlete: the direct method and the indirect method. For the direct method, state anxiety is assessed shortly before competitions until an athlete achieves a personal best performance. The optimal anxiety zone is then created by adding and subtracting 4 units to this anxiety score, which is approximately one-half standard deviation based on the STAI. The span of the optimal anxiety zone was established from initial research, but more recent studies have demonstrated that it can vary in width across athletes, indicating that some individuals can tolerate a wide range of anxiety intensity before experiencing a decline in performance. Unfortunately, the direct technique requires measuring anxiety prior to competitions

until the athlete achieves an outstanding or personal best performance, a process that may require months or even years and that is compounded when dealing with large teams consisting of many athletes.

As a more efficient alternative, Hanin developed an indirect method based on retrospection of past competitions. With the indirect method, athletes complete the state portion of the STAI according to how they recall feeling prior to their own best past performances or prior to performances judged to be optimal. Again, 4 anxiety units are added and subtracted from this "recalled best" anxiety score to establish the optimal zone. The accuracy of the indirect method has been tested by correlating recalled anxiety scores with anxiety values actually obtained at the time of the recalled events, and effect sizes ranging from .60 to .80 have been reported consistently. Research indicates that the levels of accuracy in recalling anxiety are comparable in athletes who performed either better or worse than expected despite concerns that performance outcome could bias accuracy. Although the effect size for recalled and actual precompetition anxiety supports the use of the indirect method for determining optimal anxiety ranges, occasional reports of inaccurate recall have been noted in the literature. In such instances, the direct method should be used.

3.3. Interindividual Variability in Optimal Anxiety

The IZOF model predicts not only that athletes of a similar caliber in a given sport will differ in the intensity of optimal anxiety but also that a significant proportion will benefit from high anxiety. The findings from studies of athletes in a number of sports and different levels are consistent with these propositions, indicating that between 25 and 50% report that best performances occur when anxiety levels are elevated. As predicted by the IZOF model, the proportion of athletes who perform best at high anxiety intensities is not related to skill or even age. For example, in studies of elite U.S. distance runners, it was found that 30% of female runners reported that best performances were most likely when anxiety was significantly elevated, but percentages as high as 51% have been noted in nonelite college track and field athletes. Even in track and field athletes as young as 9 to 12 years, more than one in four indicated that they performed best with high anxiety. These results and other findings provide support for the IZOF model, but they are not consistent with group-based explanations of anxiety and performance. For example, according to both

drive theory and the inverted-U hypothesis, the proportion of individuals who benefit from high anxiety should consistently be higher for elite athletes than for less talented competitors.

The evidence of wide-ranging variability in precompetition and optimal anxiety complicates the use of intervention strategies designed to regulate anxiety. Group-based interventions, in which an entire team of athletes receives a single treatment such as relaxation, are easily administered but would be ineffective or counterproductive for those athletes who perform best at moderate or high anxiety intensities. On the other hand, it can be time-consuming to use IZOF procedures to assess anxiety at the time of competition and then compare these values with each athlete's optimal zone to determine who will be in need of some form of intervention.

3.4. Predicting Precompetition Anxiety

In an effort to provide a means for coaches and psychologists to efficiently identify those athletes who will require anxiety intervention, as well as to determine the appropriate direction of the intervention (e.g., increase or decrease anxiety), IZOF studies have tested the ability of athletes to anticipate the intensity they will experience prior to actual competition. In this research, athletes completed the state portion of the STAI several days prior to a competition with instructions to respond according to how they anticipated they would feel immediately before the competition. Athletes then completed the STAI again just before the competition under the standard instructional set (i.e., "right now"), and that score was then compared with the predicted anxiety score. To minimize intrusiveness, the questionnaires were sometimes completed at a prescribed time before competition (e.g., 60 minutes). The results of this work reveal that predicted anxiety scores correlate quite closely with actual values. Correlations between predicted and actual precompetition anxiety range between .60 and .80, with higher coefficients occurring for difficult or important competitions.

Predicted precompetition anxiety scores are useful in situations where it would be difficult or impossible to assess anxiety just prior to the actual competition. The discrepancy between predicted precompetition anxiety and IZOF values can be contrasted to identify athletes who are likely to be too relaxed or too anxious, and this can be done several days before the competition.

From a practical perspective, the extent to which anxiety deviates from the optimal zone can help dictate how much anxiety must be increased or decreased to reach the optimal zone. Simple techniques that are easily implemented by the coaching staff can be an expeditious means to manipulate anxiety. These include emphasizing or deemphasizing the importance of the competition or the expectations of an athlete's performance.

3.5. Impact of Optimal Anxiety on Performance

Research based on IZOF procedures indicates that deviations in precompetition anxiety from the optimal zone have a significant impact on performance. Studies of athletes in sports such as swimming, rhythmic gymnastics, and ice skating indicate that precompetition anxiety of successful performers is closer to their own optimal anxiety values than is the case with their less successful teammates. In most cases, performance differences are evident only in difficult competitions, and it has been speculated that optimal anxiety is not necessary to achieve adequate performance in easy or unimportant competitions. In research that has examined the net impact of anxiety on sport performance, it has been found that performances were approximately 2% worse on average in cases where precompetition anxiety fell outside the IZOF and that the decrement was approximately equal whether anxiety was lower or higher than optimal.

Despite evidence that supports the major tenets of the IZOF model, the model has been criticized on a number of grounds. One primary concern is that the IZOF model does not indicate what variables contribute to the differences in optimal anxiety observed in homogenous groups of athletes. To date, studies of the factors contributing to interindividual variability in optimal anxiety have been rare, and the bulk of IZOF research has focused on examining the validity of the major tenets of the model. A second line of criticism contends that IZOF research based on the STAI or other nonspecific and general measures of anxiety is inadequate, whereas more complete results would be yielded by employing a sport-specific measure assessing multiple aspects of anxiety, particularly the CSAI-2. However, the results of CSAI-2 research on the IZOF model have been less consistent than research based on the STAI, and the IZOF model has been found to be less accurate than the STAI in assessing both recalled and predicted precompetition anxiety.

4. OTHER IDEOGRAPHIC THEORIES

Two other major theories of anxiety and sport performance have been adopted recently for sport: reversal theory and catastrophe theory. Although the concept of individual differences are not central to these theories, like the IZOF model, they acknowledge that anxiety can either facilitate or harm sport performance. Unlike the IZOF model, both reversal and catastrophe theories incorporate specific anxiety or perceived arousal scales.

Reversal theory predicts that self-reported arousal is important to performance despite a lack of evidence that self-reports can provide an objective indication of physiological activity. Arousal is interpreted on the basis of an individual's current emotional state that results from the interaction of oppositional high-arousal preferring (paratelic) and low-arousal preferring (telic) states that are also assessed via self-reports.

Catastrophe theory uses an arousal-related measure referred to as somatic anxiety, which is assumed to exhibit an inverted-U relationship with performance. In addition, measures of self-confidence and cognitive anxiety are assessed using a modified version of the CSAI-2 that assesses facilitative and debilitative anxiety. Cognitive anxiety is posited to be negatively related to performance, whereas self-confidence exhibits a positive relationship. When these dimensions are considered collectively, they form a complex three-dimensional figure referred to as a butterfly or catastrophe cusp. Reviews of the efficacy of reversal and catastrophe theories have been mixed, and this may stem in part from the challenges of validating these complex theories empirically.

5. SUMMARY

The growing realization that traditional theoretical explanations of anxiety and performance fair poorly when applied to sport has led to the development of several models and theories of anxiety specific to athletes. Further research is needed before definitive judgments can be made about the relative efficacy of these theoretical explanations, but they all indicate that influence of anxiety on sport performance is more complex than is predicted by traditional explanations such as the inverted-U hypothesis. In particular, the results of IZOF model research reveal that the anxiety intensity associated with optimal sport performance varies considerably across athletes, even for those competing in the same competition. This research also indicates that a substantial percentage of athletes actually benefit from elevated anxiety and that in these cases, interventions aimed at reducing anxiety may be counterproductive.

See Also the Following Articles

Arousal in Sport ■ Performance Slumps in Sport: Prevention and Coping ■ Psychological Skills Training in Sport

Further Reading

Duffy, E. (1957). The psychological significance of the concept of "arousal" or "activation." *Psychological Review*, 66, 183–201.

Fazey, J., & Hardy, L. (1988). *The inverted-U hypothesis: A catastrophe for sport psychology? Leeds, UK: White Line Press.*

Griffith, C. R. (1929). *The psychology of coaching.* New York: Scribner.

Hackfort, D., & Schwenkmezger, P. (1993). Anxiety. In R. M. Singer, M. Murphey, & L. K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 328–364). New York: Macmillan.

Hackfort, D., & Spielberger, C. D. (Eds.). (1989). *Anxiety in sports.* New York: Taylor & Francis.

Hanin, Y. L. (Ed.). (2000). *Emotion in sports.* Champaign, IL: Human Kinetics.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton.

Martens, R., Vealey, R. S., & Burton, D. (1990). *Competitive anxiety in sport.* Champaign, IL: Human Kinetics.

Neiss, R. (1988). Reconceptualizing arousal: Psychobiological states in motor performance. *Psychological Bulletin*, 103, 345–366.

Ostrow, A. C. (1996). *Directory of psychological tests in the sport and exercise sciences* (2nd ed.). Morgantown, WV: Fitness Information Technology.

Oxendine, J. B. (1970). Emotional arousal and motor performance. *Quest*, 13, 23–32.

Raglin, J. S., & Hanin, Y. L. (2000). Competitive anxiety. In Y. L. Hanin (Ed.), *Emotion in sports* (pp. 93–111). Champaign, IL: Human Kinetics.

Raglin, J. S. (1992). Anxiety and sport performance. In J. O. Holloszy (Ed.), *Exercise and sport sciences reviews* (Vol. 20, pp. 243–274). New York: Williams & Wilkins.

Spielberger, C. D., Gorsuch, R. L, Lushene, R. E., Vagg, P. R, & Jacobs, G. A. (1983). *Manual for the State–Trait Anxiety Inventory (Form Y).* Palo Alto, CA: Consulting Psychologists Press.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18, 459–482.

Anxiety Disorders in Late Life

Cheryl N. Carmin
University of Illinois at Chicago, Illinois, USA

Jan Mohlman
Syracuse University, Syracuse, New York, USA

Amy Buckley
University of Illinois at Chicago, Illinois, USA

GLOSSARY

cognitive behavior therapy A collaborative form of psycho-therapy that involves facilitating patients' changing their belief system via a Socratic dialogue focusing on those distressing thoughts that exacerbate negative emotions and maladaptive behaviors.

ego-syntonic/ego-dystonic obsessions Terms typically used in describing whether obsessions are congruent or incongru-ent with one's beliefs; an ego-dystonic obsession may in-clude the belief that washing one's hands for 15 minutes will prevent cancer with the understanding that, despite the persistence of the thought, it is illogical.

epidemiology A branch of medical science that deals with the incidence/prevalence, distribution, and prevention of a disease or condition as well as the identification of at-risk populations.

Likert scale A means of assigning a numerical intensity rat-ing that is anchored by descriptive terms to a construct (e.g., I would describe my worry as 1 = *very mild,* 2 = *mild,* 3 = *average,* 4 = *excessive,* 5 = *very excessive*).

Recent epidemiological studies have underscored the ubiquitous nature of anxiety disorders, with approximately 25% of adults being affected over the course of their lifetimes. Furthermore, the economic burden of these disorders has been estimated at $42.3 billion, comprising 31% of psychiatric treatment costs. Given the prevalence of anxiety disorders, it is not surprising that an increasing amount of attention has been given to investigating the prevalence and treatment of these conditions. What is surprising, however, is how little attention has been given to anxiety disorders in what is the fastest growing segment of the population, namely the elderly. This article summarizes how the existing research literature informs us with respect to the epidemiology of anxiety disorders in the elderly and then examines the treatment outcome literature with regard to the individual anxiety disorders.

1. EPIDEMIOLOGY OF ANXIETY DISORDERS IN OLDER ADULTS

The United States is aging. The U.S. Census Bureau predicts that more than 72 million adults will be age 65 years or over by the year 2030. Recent epidemiological studies underscore that anxiety disorders are the most prominent of the psychiatric conditions, with lifetime estimates nearing 25%. Despite the widespread preva-lence and extensive cost to society of anxiety disorders, there is currently little data on the rate and phenom-enology of anxiety disorders in the elderly population.

1.1. General Prevalence Data

It is generally believed that anxiety disorders occur less frequently in the elderly population than in younger adults. However, there have been only a small number of epidemiological studies that have tested this contention directly. A recent epidemiological study in the United States that included persons over 65 years of age is the Epidemiological Catchment Area (ECA) study, which included more than 18,000 noninstitutionalized adults. The ECA study systematically examined the rate of anxiety disorders at five sites across the country. Fully 30% of the ECA subjects were age 65 years or over. It was ascertained that the 1-month prevalence rate of anxiety disorders in the elderly was approximately 5.5%, lower than the 7.3% estimate for all adults surveyed. Furthermore, elderly women were nearly twice as likely to have anxiety disorders as were elderly men. The percentage of elderly adults with anxiety disorders was higher than that with any other psychiatric illness, including cognitive impairment, underscoring the need for appropriate identification and treatment in this population. Furthermore, rates of anxiety disorders were much higher than the 2.5% prevalence rate of affective disorders in this population. Other reviews of smaller scale epidemiological studies found that rates of anxiety disorders ranged between 0.7 and 19.0%. Similar 6-month prevalence rates of 3.5% have been found in Europe.

1.2. Epidemiology of Individual Anxiety Disorders

Rates for specific anxiety disorders also vary across studies and differ between elderly males and females. Data from the ECA study suggest that phobias are the most frequently experienced anxiety disorder in the elderly, affecting an estimated 4.8% of older adults. Panic disorder (PD) and obsessive–compulsive disorder (OCD) in the elderly population are expected to occur only infrequently. Rates for PD are estimated to not exceed 0.3%, and when panic does occur in elderly adults, it tends to be in women. Similarly, rates of OCD among the elderly are low, with prevalence estimates being 3.5% at a maximum and with elderly adults residing in institutional settings accounting for the upper end of these prevalence estimates.

Data regarding the prevalence of generalized anxiety disorder (GAD) suggest that it has a much more variable occurrence in the elderly population, ranging between 0.7 and 7.1%. It is important to note that data on the prevalence of GAD are not included in the ECA data and that other studies that find higher prevalence rates of GAD may be due to differences in their respective methodologies, for example, the decision rules applied in rendering a diagnosis.

1.3. Reliability of Epidemiological Studies

There are several potential explanations for the wide variability of estimates of anxiety disorder prevalence in the elderly. The existing epidemiological studies use different methodologies, and this often makes it difficult to compare prevalence rates across individual studies. Some of the studies include institutionalized elderly adults, whereas others survey community-dwelling residents only. In addition, measures of anxiety symptoms validated on younger adults might not be applicable to older adults given that the experience and expression of anxiety may change with age. Likewise, the instruments used to assess anxiety in older adults might not have norms or other psychometric data that are established for older adults, let alone for very old adults. Further research is needed to delineate factors that contribute to the difference in prevalence estimates and to clarify the nature of anxiety in the elderly. Until then, estimates of prevalence rates of anxiety disorders in the elderly should be considered preliminary.

1.4. Comorbidity

It is well known in the literature that anxiety disorders often co-occur with other diagnoses, including depression and other anxiety disorders. However, it is important to examine whether this is true for elderly persons as well. Research is beginning to shed light on the co-occurrence of anxiety and other psychiatric disorders in older adults. As with younger adults, depression most frequently co-occurs with anxiety disorders among elderly patients. In addition, when depression is the primary diagnosis in elderly adults, anxiety frequently co-occurs. In general, the delineation between depression and anxiety is not clear, and research suggests that it might be even less clear for older adults.

Unlike anxiety or depression, the incidence of cognitive impairment increases with age. Research examining the relationship between anxiety and dementia in older adults typically finds that these syndromes often coexist. Symptoms of anxiety often occur in the

context of dementia as well as in nondemented older adults.

In summary, although the frequency with which anxiety disorders are present in older adults is lower than that in younger adults, these data may be confounded by methodological problems such as differences in diagnostic classification, commonalities between anxiety and depression in this population, and the frequent occurrence of anxiety symptoms accompanying medical conditions.

2. ASSESSMENT AND TREATMENT OF ANXIETY DISORDERS

2.1. Generalized Anxiety Disorder

2.1.1. Assessment

Diagnostic criteria for GAD have undergone considerable revision over the past 20 years or so. The most recent version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) emphasizes chronic uncontrollable worry that causes significant distress or impairment as the hallmark symptom of the disorder. In adults of all ages, GAD is most often comorbid with other affective and anxiety disorders and is sometimes viewed as a vulnerability factor for the development of additional psychiatric problems. GAD is often secondary to depressive disorders in older adults.

As noted by Carmin and colleagues in 1999, the only clinician-rated diagnostic interviews that have been found to yield reliable diagnoses with elderly patients are the Structured Clinical Interview Diagnostic (SCID) and the Anxiety Disorders Interview Schedule (ADIS-IV). Low interrater reliability estimates were reported during the 1980s for GAD; however, better estimates have been found using DSM-IV criteria. A recent investigation using the SCID and the ADIS yielded substantial diagnostic agreement for all of the anxiety disorders in an older treatment-seeking sample. Interrater reliability for GAD with the ADIS was similarly found to be excellent.

Two additional clinician-rated measures have been used to determine anxiety severity in the elderly. The Hamilton Anxiety Rating Scale (HARS) is a 14-item scale tapping anxiety symptoms in somatic, psychic, and affective domains. In 1997, Beck and Stanley found that the HARS distinguished older adults with GAD from normal controls, providing some preliminary support for the measure. The FEAR is a relatively new instrument developed to measure GAD in older adults in the primary care setting. The FEAR, a 4-item version of the 11-item Anxiety Disorder Scale, is meant to be administered verbally during routine medical exams or in the waiting room. The FEAR demonstrated very good sensitivity and specificity in an initial administration to 88 older medical patients, 27% of whom were diagnosed with GAD. Additional psychometric properties of the measure await investigation.

There are only a few self-report measures that are recommended for use with late-life GAD patients. The Worry Scale (WS) is a 35-item, self-report questionnaire tapping three subscales of worry domains of importance to older adults—finances, health, social issues—using a 5-point Likert scale. The primary utility of the measure is the assessment of GAD. However, the use of the measure with a sample of older adults with GAD yielded poor convergent validity, with low correlations with other measures of anxiety. The internal reliability estimate was .93, but test–retest reliability was not reported. In a second study by Beck and Stanley in 1997, the measure showed adequate internal reliability and convergent validity. Currently, the properties of this measure seem to be better established in healthy controls than in patient samples.

The Penn State Worry Questionnaire (PSWQ) consists of 16 items on a 4-point Likert scale designed to tap the generality, intensity, and uncontrollability of worry. The measure has demonstrated good psychometric properties (e.g., internal consistency, convergent validity) in several studies of older adults with GAD. However, the PSWQ failed to discriminate between older adults with GAD and PD and has shown inadequate test–retest reliability in older GAD patients over a variable period with a mean of 70 days. Stanley and colleagues found that the PSWQ showed good divergent validity with measures of depression. Because late-life GAD is also comorbid with depressive disorders, this is one appealing feature of the measure.

The trait scale of the State–Trait Anxiety Inventory (STAI) assesses the tendency to experience frequent anxiety and nervousness. The measure has 20 items on a 4-point Likert scale and is one of the most frequently used self-report questionnaires in studies of anxiety. In heterogeneous older adult samples, the trait scale has shown adequate psychometric properties, and there are ample normative data. In 1996, Stanley, Beck, and Zebb found low to moderate correlations between the STAI and other anxiety and worry measures in older GAD patients but found good internal consistency of .88 for the trait scale. Test–retest

reliability estimates were not given for the GAD sample but were good in a sample of healthy older adults. A more recent investigation showed good internal consistency but low test–retest reliability (.58) in a sample of 57 older adults with GAD. Mohlman, deJesus, and colleagues found that scores on the measure were very similar in older adults with GAD and PD, suggesting that it taps a nonspecific construct in older samples.

2.1.2. Psychosocial Treatment of Late-Life Generalized Anxiety Disorder

Two early studies by Stanley and colleagues indicated that supportive group therapy was beneficial for late-life GAD treatment. The 14-week intervention focused on the discussion of symptoms and experiences and providing support for group members.

Investigations of individual format cognitive behavior therapy (CBT) have suggested that this treatment is effective in older adults with GAD. However, it is notable that most of these studies have included nonstandard augmentations to the therapy such as concurrent medication and weekly meetings with a physician, treatment conducted in primary care or in patients' own homes, and the use of learning aids. The only investigation of standard individual CBT delivered in a mental health clinic indicated very modest efficacy as compared with a wait list condition. Currently, the efficacy of individual CBT is not well supported.

Several studies assessing the efficacy of group CBT for late-life GAD indicate that the treatment is typically more effective than wait list control conditions; however, it is not significantly better than other control conditions. Thus, it is possible that nonspecific elements of the group format (e.g., increased social interaction, mitigation of loneliness), rather than elements specific to CBT, led to improvement. Studies comparing group format CBT with individual format CBT should help to clarify this issue.

2.2. Panic Disorder and Agoraphobia

2.2.1. Assessment

Notably, very few studies have addressed the psychometric properties of measures of panic and related symptoms in older patient samples. Two studies, one using a nonclinical community sample and the other using patients drawn from a medical clinic, studied the properties of the Beck Anxiety Inventory (BAI), a well-known

measure of anxiety and panic symptoms and found that the scale had good discriminant validity and internal consistency. The latter study also demonstrated the BAI four-factor solution with autonomic, neuromotor, cognitive, and panic subscales, suggesting that anxiety symptom clusters are slightly different from those found among younger adults.

The Anxiety Sensitivity Index (ASI) is a 16-item measure tapping the fear of anxiety sensations, which is known to be a risk factor for the development of panic. In 2000, Mohlman and Zinbarg tested the structure and validity of the ASI in 322 healthy older adults (mean age 75 years). The ASI showed strong internal consistency and moderate correlations with measures of related constructs. Confirmatory factor analysis indicated a hierarchical structure with three group factors—physical concerns, mental incapacitation concerns, and social concerns—as well as a general factor, consistent with previous investigations of the ASI in younger adults. In 1998, Deer and Calamari found that 49% of their older sample (mean age 81 years) reported panic symptoms and that 27% reported a panic attack during the past year. Anxiety sensations predicted unique variance in panic symptomatology and may function as a risk factor for the development of late-life panic.

2.2.2. Psychosocial Treatment of Late-Life Panic Disorder

Trials of psychosocial treatments in older samples with PD, panic disorder with agoraphobia (PDA), or agoraphobia without history of panic disorder (AWOHPD) are limited to case studies and three small pilot studies. One study found that principles of reality therapy, which focuses on an individual's situation and worldview, were effective when used by a neighbor to mitigate an older adult's paranoia and agoraphobia.

Early investigations of behavioral treatments conducted during the 1970s included relaxation, imagery, and exposure. In 1996, Rathus and Sanderson used CBT with two older panic patients: one 70-year-old male and one 69-year-old female. Treatment components were education, cognitive restructuring, interoceptive and situational exposure, and diaphragmatic breathing. Both participants achieved panic-free status and decreased depression following 4 to 5 months of therapy.

In 1991, King and Barrowclough tested CBT for panic and anxiety in a small sample of adults ages 66 to 78 years. Of the 10 participants, 8 had primary diagnoses of PDA. After treatment, 7 were free of

panic and 2 showed decreased symptom severity. Six months later, 8 of the remaining 9 participants were panic free and 6 showed improvement on depression.

In 1996, Swales and colleagues tested 10 90-minute sessions of CBT in 15 adults ages 55 to 80 years. Participants experienced decreased severity and frequency of panic attacks, depression, avoidance, and role impairment, and this was apparent at both posttreatment and 3-month follow-up. A reanalysis of Gorenstein and colleagues by Mohlman indicated that CBT plus medication management ($n = 5$) was somewhat more effective than medication management alone ($n = 5$) in assisting older adults with PD to decrease anxiety while tapering off anxiolytic medication. The use of interoceptive exposure was believed to facilitate habituation to sensations related to PD and medication withdrawal simultaneously.

2.3. Social Anxiety Disorder

Although the social anxiety literature has grown tremendously during the past several years, there is still a relative dearth of empirical research examining the assessment and treatment of social anxiety in older adults. This may be due in part to the fact that in epidemiological studies, social phobia is a relatively rare disorder among the elderly population. In addition, social anxiety as a distinct diagnostic category did not appear until the third edition of the DSM was published.

2.3.1. Assessment

As noted previously, the SCID and the ADIS-IV are the only two clinician-administered diagnostic instruments with published data on older adults. In 1993, Segal and colleagues reported interrater reliabilities ranging from good to excellent across anxiety disorder diagnoses with use of the SCID.

The ADIS-IV is considered the gold standard for diagnosing anxiety disorders in adults. In 2001, Brown and colleagues found the interrater reliability of the lifetime version of this instrument to be excellent for social anxiety on a population of younger adults. Its reliability for diagnosing social phobia in older adults has not yet been established.

There are currently a number of self-report measures available for assessing social anxiety in younger adults. Unfortunately, none of the measures that has been used with younger cohorts has norms or other psychometric data supporting its use with older adults.

2.3.2. Treatment

In contrast to the abundant literature examining psychological and pharmacological treatment of social anxiety disorder in adults, there are no published studies examining the efficacy or effectiveness of treatments for socially anxious older adults. In 2001, Fedoroff and Taylor conducted a meta-analysis of 108 studies. These researchers compared pharmacological treatments, including selective serotonin reuptake inhibitors (SSRIs), benzodiazepines (BZDs), and monoamine oxidase inhibitors (MAOIs), with components of CBT, including exposure and cognitive restructuring. Studies were included in the analysis if they consisted of at least four patients diagnosed with social anxiety using clinical interviews and consisted of standard treatments of social anxiety. The ages of the patients included in the meta-analysis were not provided, although all studies consisted of adults. Results of the analysis suggested that pharmacotherapies were superior to CBT at posttreatment and that both were superior to controls. BZDs were significantly more effective than both CBT and controls but did not significantly differ from SSRIs. Long-term maintenance of gains was not as well established because many of the studies did not contain sufficient information for estimating this variable. Although these results may have applicability to older adults, the use of BZDs is ventured very cautiously given that this class of medication can cause difficulties with dizziness, respiration, and cognitive functioning.

In summary, the literature is scarce regarding effective treatments for social anxiety in older adults. Although both CBT and pharmacotherapy have been proven to be effective treatments, extrapolation to older adults is premature. More research is needed in this area before conclusions can be made concerning best practice treatments of social anxiety treatment in older adults.

2.4. Assessment and Treatment of Obsessive–Compulsive Disorder

The observation made by McCarthy and colleagues more than a decade ago, that there is a paucity of research evaluating the efficacy of treatments for older adults diagnosed with OCD, still remains true today. The limited literature that exists has focused primarily either on a variation of CBT called exposure and response/ritual prevention (ERP) or on pharmacological treatment.

2.4.1. Assessment

The assessment literature, with regard to OCD in older adults, offers few recommendations other than the Padua Inventory, which was used with older adults diagnosed with GAD. Some authors have suggested that certain presentations of OCD, such as obsessions and compulsions related to fear of forgetting names and pronounced ego-syntonic scrupulosity, are more likely to occur in the elderly. Besides these case reports, there is little evidence that particular constellations of obsessions and compulsions are unique to older adults. Typical presentations of OCD, such as contamination fears with washing rituals and fears of harming others accompanied by checking compulsions, are also commonly found in the elderly. DSM-IV criteria appear to be appropriate for use with older adults, but more research is needed using structured diagnostic instruments such as the ADIS-IV to better characterize the presentation of OCD in the elderly.

2.4.2. Treatment

Two early uncontrolled case studies of elderly OCD patients reported significant reductions in OCD symptoms following ERP, whereas one anecdotal report documented unsuccessful combined ERP and medication treatments in a 74-year-old woman with OCD and a learning disability. Several controlled outcome studies of ERP using single-case designs yielded similarly successful results. The one study that directly compared treatment responses of older patients with those of younger patients examined the effectiveness of inpatient ERP administered to 10 severely impaired OCD patients age 60 years or over and 10 younger OCD patients matched for gender and clinical severity. No significant differences in response to treatment were found between the older adults and their younger counterparts, with the majorities of both groups being classified as treatment responders at posttreatment. This finding is particularly noteworthy given that the older patients reported having been symptomatic for more than twice as long as the younger adult cohort.

In a recent controlled case comparison, Carmin and Wiegartz described two older men with OCD. One experienced a successful outcome and the other experienced an unsuccessful outcome when intensive inpatient ERP was the treatment modality. These authors concluded that the duration of the illness, comorbidity of other psychiatric disorders as well as medical conditions, and the availability of social support can have an effect on treatment outcome.

Cognitive decline can exacerbate or mimic symptoms of OCD, and medical difficulties, such as cerebrovascular accidents (e.g., basal ganglia infarcts), are more prevalent in the elderly and have been noted to produce OCD symptoms in previously healthy patients. Such observations raise the question of whether individuals who have experienced neurological insults that result in OCD can benefit from psychological treatment. Of considerable importance is that ERP and pharmacological treatment were found to be effective in treating a 65-year-old man whose OCD was related to recent basal ganglia infarcts.

In comparison with the previously noted studies that used ERP, a follow-up study of medication and supportive psychotherapy that was offered to residents of an old-age home suggested that these methods can have a positive effect on treatment of anxiety disorders, including OCD and panic. One limitation of this study was that this was a diagnostically heterogeneous group, and no details were provided about what medications were used or what supportive psychotherapy entailed.

Studies focusing on CBT for late-life OCD have consisted of relatively small patient samples, thereby limiting generalizability. Preliminary findings, however, suggest that late-life OCD is treatable. Even if subsequent research finds lower success rates for the elderly than rates typically reported for the general adult population, there is sufficient evidence to conclude that at least some, if not most, older adults respond to ERP.

2.5. Assessment and Treatment of Posttraumatic Stress Disorder

Much of the literature pertaining to posttraumatic stress disorder (PTSD) in the elderly focuses on holocaust and natural disaster victims and combat veterans. Given the vulnerability of older adults to physical violence, there is limited research that examines symptom presentations of elderly crime victims and treatment.

Typically, PTSD symptoms in older adults reflect a chronic waxing and waning of symptoms, with exacerbations linked to the expected stressors of advancing age. Despite early studies suggesting a level of resiliency in older adult disaster victims, this resiliency may reflect an underreporting of symptoms by PTSD sufferers, an underdiagnosis by clinicians, or an attribution of anxiety-related somatic symptoms to the normal frailties associated with old age by clinicians, thereby making the accurate diagnosis of this disorder difficult.

2.5.1. Assessment

There are several measures that have been used to assess PTSD symptoms in the elderly. The Clinician-Administered PTSD Scale (CAPS) is a clinician-administered semistructured interview that has been used extensively with older individuals who have typically been exposed to war-related trauma. In addition, the CAPS has been used as a process and outcome measure for those experiencing non-combat-related PTSD. These studies suggest that the CAPS is recommended for use with older PTSD sufferers.

Although a number of self-report measures have been used to assess PTSD in older adult samples, these studies typically provide descriptive rather than psychometric data. Although the Impact of Events Scale (IES) has provided psychometric data, the findings are equivocal with respect to its use with elders, suggesting that the type of traumatic event may be more important than symptomatology in older samples than in younger samples. The combat and civilian forms of the Mississippi PTSD Scales (MISS) have been significantly correlated with diagnostic measures of PTSD; however, the combat version appears to be the most accurate measure of PTSD severity in a small sample of elderly former prisoners of war.

2.5.2. Treatment

There are two reports of PTSD treatment in older adults. In 1998, Bonwick described a 16-week group day hospital treatment for veterans. The program included elements of psychoeducation, symptom management, relaxation, group therapy, and physical exercise. No outcome data were reported, but the author noted that those receiving treatment reported a greater understanding of PTSD, improved coping skills, and an enhanced quality of life. A recent conceptual review of PTSD in older adults indicates that a better understanding of the issues related to risk and vulnerability to trauma, such as the availability of social support networks, the use of coping strategies, and perceptions of the meaning related to the traumatic event, may allow for a better understanding of how to construct better treatment interventions.

2.6. Assessment and Treatment of Specific Phobias

2.6.1. Assessment

There is surprisingly little information with respect to the assessment and treatment of specific fears in older

adults given that specific phobias is the most prevalent anxiety disorder in this age group. Of the measures available that assess for the gamut of phobias, the Fear Survey Schedules (FSS-II and FSS-III) have been used with older adults and are promising screening measures for identifying specific fears. Kogan and Edelstein revised this measure specifically for use with older adults (FSS-OA). Their preliminary results are encouraging but not conclusive. One measure, the Falls Efficacy Scale (FES), appears to have good reliability and validity. However, additional psychometric evaluations (e.g., convergent and discriminatory validity) still need to be done before this self-report measure can be used independently of a comprehensive anxiety assessment battery in older persons with somatic symptoms of dizziness or balance disturbances.

2.6.2. Treatment

There is one randomized controlled study comparing CBT with an educational control group in more than 400 individuals (average age of 77 years) experiencing fears of falling. The CBT patients demonstrated posttreatment gains in mobility control and increased activity, but these gains were lost over a 6-month follow-up period. At 1 year follow-up, the CBT patients showed increased improvement in different areas (e.g., mobility range, social functioning). The authors of this study noted that only 63.4% of their patients attended more than five of the eight offered treatment sessions, thereby highlighting problems with compliance and attrition.

2.7. Issues Related to Anxiety Secondary to a Medical Condition

There is considerable overlap between many symptoms diagnostic of an anxiety disorder and symptoms that can be attributed to a medical illness. Medical illnesses such as cardiovascular disease, pulmonary dysfunction, stroke, hyperthyroidism, sensory impairments, and dementia can mimic, exacerbate, antedate, and/or accompany symptoms of anxiety. For example, symptoms of panic may overlap with certain symptoms related to angina, congestive heart failure, or emphysema, causing the diagnosis of panic to be overlooked. Alternatively, the normal developmental changes associated with aging can be mistaken for an anxiety disorder. Individuals who have sensory or mobility impairments may repeatedly check for where a hearing

aid or walker is located or ask for frequent reassurance, resulting in clinicians mistakenly suggesting a diagnosis of OCD. Clearly, differentiating between anxiety disorders and medical illness in older adults is a complicated task. There are no studies that address the treatment of comorbid anxiety and medical illness.

3. CONCLUSIONS

Enormous strides have been made in the area of the psychopathology, assessment, and treatment of anxiety disorders in adults. Unfortunately, in the area of geriatric anxiety, a tremendous amount of research still needs to be done.

There are significant questions as to whether the fundamental nature of anxiety disorders in adults is the same as that during their later years. A further complication arises in that neurobiological changes across the life span were found to cause an age-related decreased cortisol response to an experimental stressor. These findings suggest a decrease in reactivity to stress with advancing years. It has likewise been hypothesized that age-related changes in hypothalamic–pituitary–adrenal (HPA) activity, as indicated by cortisol level, are markers for central nervous system dysfunction. If this is indeed true, anxiety may be a link between central nervous system instability and the increases in cognitive impairment that are often found in aging and may explain the decreased prevalence of anxiety disorders in older adults.

As noted previously, epidemiological studies are confounded by their method of sampling. Cohort effects relevant to the stigma attached to mental illness have a greater influence on older adults. Thus, fewer elderly individuals may be willing to endorse anxiety symptoms in the course of an epidemiological study. Likewise, where samples are drawn from may influence prevalence data.

Finally, it would appear that both psychosocial and pharmacological treatments appear to aid in the reduction of anxiety symptoms in elderly samples. The prevailing form of psychotherapy that has been studied has been CBT. However, there is not firmly conclusive data that would allow one to unequivocally endorse CBT or a particular medication for use in treating a given anxiety disorder. Clearly, far more research is needed. One optimistic note is that the participants in existing longitudinal studies, such as the Harvard/Brown Anxiety Research Program (HARP), are aging. Data such as those generated by this study will allow for the close examination of how anxiety disorders progress over adulthood and into later life and, hopefully, will provide answers to many of the questions that remain regarding geriatric anxiety.

See Also the Following Articles

Depression in Late Life ■ Panic ■ Personality and Emotion in Late Life ■ Posttraumatic Disorders ■ Psychotherapy in Older Adults ■ Stress

Further Reading

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Averill, P. M., & Beck, J. G. (2000). Posttraumatic stress disorder in older adults: A conceptual review. *Journal of Anxiety Disorders*, *14*, 133–156.

Ballenger, J. C., Davidson, J. R., Lecrubier, Y., Nutt, D. J., Borkovec, T. D., Rickels, K., Stein, D. J., & Wittchen, H. U. (2001). Consensus statement on generalized anxiety disorder from the International Consensus Group on Depression and Anxiety. *Journal of Clinical Psychiatry*, *62*(Suppl. 11), 53–58.

Barrowclough, C., King, P., Colville, J., Russell, E., Burns, A., & Tarrier, N. (2001). A randomized trial of the effectiveness of cognitive–behavioral therapy and supportive counseling for anxiety symptoms in older adults. *Journal of Consulting and Clinical Psychology*, *69*, 756–762.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*, 893–897.

Beck, J. G., & Stanley, M. A. (1997). Anxiety disorders in the elderly: The emerging role of behavior therapy. *Behavior Therapy*, *28*, 83–100.

Beck, J. G., Stanley, M. A., & Zebb, B. J. (1996). Characteristics of generalized anxiety disorder in older adults: A descriptive study. *Behaviour Research and Therapy*, *34*, 225–235.

Bonwick, R. (1998). Group treatment programme for elderly war veterans with PTSD [letter]. *International Journal of Geriatric Psychiatry*, *13*, 64–65.

Brown, T. A., DiNardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of DSM-IV anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology*, *110*, 49–58.

Carmin, C. N., Pollard, C. A., & Gillock, K. L. (1999). Assessment of anxiety disorders in the elderly. In P. A. Lichtenberg (Ed.), *Handbook of assessment in clinical gerontology* (pp. 59–90). New York: John Wiley.

Carmin, C. N., Pollard, C. A., & Ownby, R. L. (1998). Obsessive–compulsive disorder: Cognitive behavioral

treatment of older versus younger adults. *The Clinical Gerontologist*, *19*(2), 77–81.

Carmin, C. N., & Wiegartz, P. S. (2000). Successful and unsuccessful treatment of obsessive–compulsive disorder in older adults. *Journal of Contemporary Psychotherapy*, *30*, 181–193.

Carmin, C. N., Wiegartz, P. S., & Scher, C. (2000). Anxiety disorders in the elderly. *Current Psychiatry Reports*, *2*, 13–19.

Carmin, C. N., Wiegartz, P. S., Yunus, U., & Gillock, K. L. (2002). Treatment of late-onset OCD following basal ganglia infarct. *Depression & Anxiety*, *15*(2), 87–90.

Deer, T. M., & Calamari, J. E. (1998). Panic symptomatology and anxiety sensitivity in older adults. *Journal of Behavior Therapy and Experimental Psychiatry*, *29*, 303–316.

Devanand, D. P. (2002). Comorbid psychiatric disorders in late life depression. *Biological Psychiatry*, *52*, 236–242.

Fedoroff, I. C., & Taylor, S. (2001). Psychological and pharmacological treatments of social phobia: A meta-analysis. *Journal of Clinical Psychopharmacology*, *21*, 311–324.

First, M., Spitzer, R., Gibbon, M., et al. (1995). *The Structured Clinical Interview for DSM-IV Axis I disorders (SCID-I, version 2.0, final version)*. New York: New York State Psychiatric Institute.

Flint, A. (1994). Epidemiology and comorbidity of anxiety disorders in the elderly. *American Journal of Psychiatry*, *151*, 640–649.

Flint, A. J., Koszycki, D., Vaccarino, F. J., Cadieux, A., Boulenger, J-P., & Bradwejn, J. (1998). Effect of aging on cholecystokinin-induced panic. *American Journal of Psychiatry*, *155*, 283–285.

Forsell, Y., & Winblad, B. (1998). Feelings of anxiety and associated variables in a very elderly population. *International Journal of Geriatric Psychiatry*, *13*, 454–458.

Gray, J. J., & Acierno, R. (2002). Symptom presentations of older adult crime victims: Description of a clinical sample. *Journal of Anxiety Disorders*, *16*, 299–309.

Kabacoff, R. I., Segal, D. L, Hersen, M., & Van Hasselt, V. B. (1997). Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State–Trait Anxiety Inventory with older adult psychiatric outpatients. *Journal of Anxiety Disorders*, *11*, 33–47.

King, P., & Barrowclough, C. (1991). A clinical pilot study of cognitive–behavioral therapy for anxiety disorders in the elderly. *Behavioural Psychotherapy*, *19*, 337–345.

Kogan, J. N., & Edelstein, B. A. (in press). Modification and psychometric examination of a self-report measure of fear in older adults. *Journal of Anxiety Disorders*.

Kogan, J. N., Edelstein, B. A., & McKee, D. R. (2000). Assessment of anxiety in older adults: Current status. *Journal of Anxiety Disorders*, *14*, 109–132.

Krasucki, C., Ryan, P., Ertan, T., Howard, R., Lindesay, J., & Mann, A. (1999). The FEAR: A rapid screening instrument for generalized anxiety in elderly primary care attenders. *International Journal of Geriatric Psychiatry*, *14*, 60–68.

Lindesay, J. (1996). Elderly people and crime. *Reviews in Clinical Gerontology*, *6*, 199–204.

Manela, M., Katona, C., & Livingston, G. (1996). How common are the anxiety disorders in old age? *International Journal of Geriatric Psychiatry*, *11*, 65–70.

McCarthy, P. R., Katz, I. R., & Foa, E. B. (1991). Cognitive–behavioral treatment of anxiety in the elderly: A proposed model. In C. Salzman, & B. D. Leibowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 197–214). New York: Springer.

Meeks, S., Woodruff-Borden, J., & Depp, C. A. (2003). Structural differentiation of self-reported depression and anxiety in late life. *Journal of Anxiety Disorders,* *17*, 627–646.

Mohlman, J., Gorenstein, E. E., Kleber, M., deJesus, M. B., Gorman, J. M., & Papp, L. A. (2003). Standard and enhanced cognitive behavior therapy for late life generalized anxiety disorder: Two pilot investigations. *American Journal of Geriatric Psychiatry*, *11*, 24–32.

Mohlman, J., deJesus, M., Gorenstein, E. E., Kleber, M., Gorman, J. M., & Papp, L. A. (in press). Distinguishing generalized anxiety disorder, panic disorder, and mixed anxiety states in older treatment-seeking adults. *Journal of Anxiety Disorders*.

Mohlman, J., & Zinbarg, R. E. (2000). The structure and correlates of anxiety sensitivity in older adults. *Psychological Assessment*, *12*, 440–446.

Ownby, R. L., Harwood, G. D., Barker, W. W., & Duara, R. (2000). Predictors of anxiety in patients with Alzheimer's disease. *Depression and Anxiety*, *11*, 38–42.

Radley, M., Redston, C., Bates, F., & Pontefract, M. (1997). Effectiveness of group anxiety management with elderly clients of a community psychogeriatric team. *International Journal of Geriatric Psychiatry*, *12*, 79–84.

Rathus, J. H., & Sanderson, W. C. (1996). Cognitive behavioral treatment of panic disorder in elderly adults: Two case studies. *Journal of Cognitive Psychotherapy*, *10*, 271–280.

Schaub, R. T., & Linden, M. (2000). Anxiety and anxiety disorders in the old and very old: Results from the Berlin Aging Study (BASE). *Comprehensive Psychiatry*, *41*(2), 48–54.

Segal, D. L., Hersen, M., VanHasselt, V. B., Kabacoff, R. L., & Roth, L. (1993). Reliability of diagnosis in older psychiatric patients using the Structured Clinical Interview for DSM-III-R. *Journal of Psychopathology and Behavioral Assessment*, *15*, 347–356.

Sheikh, J. I., & Cassidy, E. L. (2000). Treatment of anxiety disorders in the elderly: Issues and strategies. *Journal of Anxiety Disorders*, *14*, 173–190.

Snyder, A. G., Stanley, M. A., Novy, D. M., Averill, P. M., & Beck, J. G. (2000). Measures of depression in older adults with generalized anxiety disorder: A psychometric evaluation. *Depression and Anxiety*, *11*, 114–120.

Spielberger, C., Gorsuch, R., Lushene, R., Vagg, P., & Jacobs, G. (1983). *State–Trait Anxiety Inventory: Test manual.* Palo Alto, CA: Consulting Psychologists Press.

Stanley, M. A., & Beck, J. G. (2000). Anxiety disorders. *Clinical Psychology Review*, *20*, 731–754.

Stanley, M. A., Beck, J. G., Barron, M., & Wagner, A. L. (1994, November). *Treatment of generalized anxiety disorder in older adults: Examination of outcome predictors.* Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, San Diego.

Stanley, M. A., Beck, J. G., & Glassco, J. D. (1996). Treatment of generalized anxiety in older adults: A preliminary comparison of cognitive–behavioral and supportive approaches. *Behavior Therapy*, *27*, 565–581.

Stanley, M. A., Beck, J. G., Novy, D. M., Averill, P. M., Swann, A. C., Diefenbach, G. J., & Hopko, D. (2003). Cognitive behavioral treatment of late-life generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 71*, 309–319.

Stanley, M. A., Beck, J. G., & Zebb, B. J. (1996). Psychometric properties of four anxiety measures in older adults. *Behaviour Research and Therapy*, *34*, 827–838.

Stanley, M. A., Novy, D. M., Bourland, S. L., Beck, J. G., & Averill, P. M. (2001). Assessing older adults with generalized anxiety: A replication and extension. *Behaviour Research and Therapy*, *39*, 221–235.

Swales, P., Solfrin, J., & Sheikh, J. (1996). Cognitive behaviour therapy in older panic disorder patients. *American Journal of Geriatric Psychiatry*, *4*, 46–60.

Swartz, M., Martin, T., Martin, M., Eizur, A., & Barak, Y. (1999). Outcome of a psychogeriatric intervention in an old-age home: A 3 year follow-up study. *Annals of Clinical Psychiatry*, *11*, 109–112.

Van Balkom, A. J., Beekman, A. T., De Beurs, E., Deeg, D. J., Van Dyck, R., & Van Tilburg, W. (2000). Comorbidity of the anxiety disorders in a community-based older population in The Netherlands. *Acta Psychiatrica Scandinavica*, *101*, 37–45.

Wetherell, J. L., Gatz, M., & Craske, M. G. (2003). Treatment of generalized anxiety disorder in older adults. *Journal of Consulting and Clinical Psychology*, *71*, 31–40.

Arousal in Sport

Leonard D. Zaichkowsky and Adam H. Naylor

Boston University, Boston, Massachusetts, USA

GLOSSARY

anxiety A negative emotional state generally expressed as worry and apprehension and accompanied by heightened arousal.

arousal The body's degree of activation from low to high; it is reflected in hormone levels and other physiological reactivity that is mediated by thoughts and feelings.

stress The body's reaction to perceived threat by an individual; the reaction is typically increased anxiety and arousal.

Arousal in sport and exercise is a human condition that ranges along a continuum from sleep to high excitation and is expressed physiologically, cognitively, and behaviorally. Three prominent theories of arousal in sport are drive theory, the inverted-U hypothesis, and reversal theory. Each has unique insights for the discussion of an athlete's ability to perform at optimal levels on the playing field or court and can guide an astute coach, counselor, or psychologist in his or her work. Lastly, the application of cognitive, behavioral, and cognitive–behavioral interventions for arousal regulation are discussed.

1. INTRODUCTION

The coach gathers the team moments before the championship game for a pep talk laden with fire, brimstone, and proclamations of victory. The professional basketball player sits on the training room table, tattoos lining his arms, bobbing his head rhythmically to the music on his headphones. The sprinter takes a few quick strides, bounds up and down a couple of times, and takes two good deep breaths before assuming her position in the starting blocks. The stock car racer says a short prayer and finds a brief moment of silence before sliding into his car and heading out to the starting line. Excellent athletic accomplishments require the athlete's body to be appropriately energized, with physiological and psychological resources prepared for the stresses and physical demands of competition.

The preceding instances describe the concept of arousal and its regulation in sport. The term "arousal" has been a part of the English language for more than 400 years. The *Oxford English Dictionary* defines "arouse" as "to raise up from sleep or inactivity, to excite." James, in his classical work *Principles of Psychology,* used the term "arousal" to describe "activation." Arousal, as it is used in sport, describes a condition of controlled or uncontrolled preparedness and

manifests itself in varying levels of mental activity, emotional display, physical energy, and concomitant physiological reactivity.

2. CONCEPTUAL CONFUSION SURROUNDING AROUSAL

Besides the historical reference of exciting one to action, being aroused typically refers to physiological responses. That is, aroused individuals demonstrate high heart rates, increased sweat response, increased brain activity, and (typically) increased intensity of physical effort.

A perusal of the general psychology and sport psychology research over the past five decades shows that arousal is often used interchangeably with terms such as anxiety, activation, emotion, motivation, and psychic energy. This understandably has resulted in conceptual confusion. The early researchers of arousal, such as Cannon, Duffy, Hebb, and Malmo, treated arousal as a unidimensional physiological construct. Today, arousal is viewed as a multidimensional construct that includes a physiological dimension paired/grouped with cognitive, affective, and/or behavioral dimensions. For example, an ice hockey player who is optimally aroused would have an elevated heart rate, increased beta waves, increased respiration, and increased adrenalin. In addition, thought processes would be alert and focused, such that the player is able to read and react to on-ice situations quickly and accurately. The optimally aroused athlete would also be in control of essential affective states and would demonstrate "passion" or intensity in play.

The physiological component of arousal can be represented in a number of ways such as through muscle tension, cortical activity, electrodermal activity, respiration, and biochemical markers (e.g., epinephrine, cortisol). The cognitive interpretation appraisal component can influence the physiological component. Furthermore, the affective component can also interact with thoughts to influence physiological responses and athletic performance (Fig. 1).

3. NEUROPSYCHOLOGY OF AROUSAL

During recent years, much has been learned about the neurological and biochemical mechanisms involved in human arousal. Preparation for competition, as well as competition itself, sets off a chain of events in the central nervous system and autonomic nervous system that results in arousal. Table I provides a summary of nervous system structures involved in arousal as well as the locations and functions of the structures.

4. AROUSAL THEORIES

A number of theories have been proposed to explain the relationship between arousal and athletic performance. Following is an analysis of the major perspectives.

4.1. Drive Theory

The drive theory of arousal states that arousal has a positive correlation with athletic performance, particularly in cases of skills that are well learned (Fig. 2). The theory is based on the classical work of Hull, who stated that drive (or arousal) multiplied by habit strength (amount of prior learning) determines an individual's level of performance:

$$\text{Performance} = \text{Drive} \times \text{Habit Strength}.$$

As emotions and excitement increase, so does an athlete's performance.

Mediating the role of anxiety on an athlete's performance is habit strength. Habit strength speaks to how well an individual has learned particular skills. The better learned a skill (or the more skilled the athlete), the more likely arousal is to enhance athletic performance. Some researchers have even gone so far as to suggest that the play of novice athletes (low in habit strength) is inhibited by increasing arousal levels.

Although this theory has many merits and has been closely associated with the concept of social facilitation, some have questioned whether the arousal–performance relationship depicted exists in all motor activities required in sport. The hypothesis is
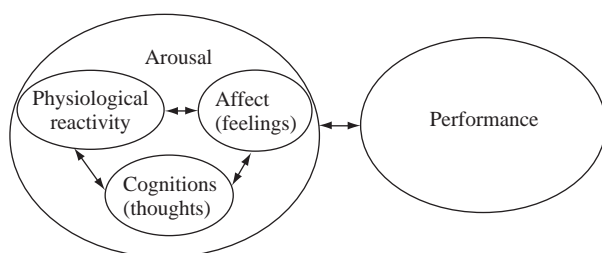
FIGURE 1   Arousal and performance.

Structure	Location	Function
Reticular formation	Part of brain stem that continues out of medulla and pons	Involved in sleep, waking, alertness, and optimal brain arousal
Hypothalamus	Between the thalmus and midbrain	Involved in appetitive and sexual behavior, regulation of sympathetic nervous system and pituitary gland
Limbic system	Series of structures, including the hypocampus and amygdala, located near the border between the cerebral hemisphere and the brain stem; the hypothalmus is sometimes considered part of the limbic system	Primarily concerned with emotional behavior; amygdala appears to be involved in aggression
Cerebral cortex	Convoluted outer layer of the human brain	Highest center involved in learning, remembering, planning, and performing motor acts
Sympathetic nervous system	A branch of the autonomic nervous system with nerve fibers originating in thoracic and lumbar regions of the spinal cord	Activates glands and smooth muscles during arousal
Parasympathetic nervous system	A branch of the autonomic nervous system with nerve fibers originating in the brainstem and sacral regions of the spinal cord	Maintains appropriate internal states during times of relaxation
Adrenal cortex	Outer layer of two small endocrine glands just above the kidneys	Secretes cortisol that regulates metabolism and stress response
Adrenal medulla	Inner layer of two small endocrine glands just above the kidneys	Secretes adrenaline (epinephrine) and noradrenalin (norepinephrine); both are involved in increased activation of the body
Pituitary gland	Deep inside brain, just below hypothalmus	Releases adrenocorticotropic hormone that stimulates adrenal cortex to release cortisol
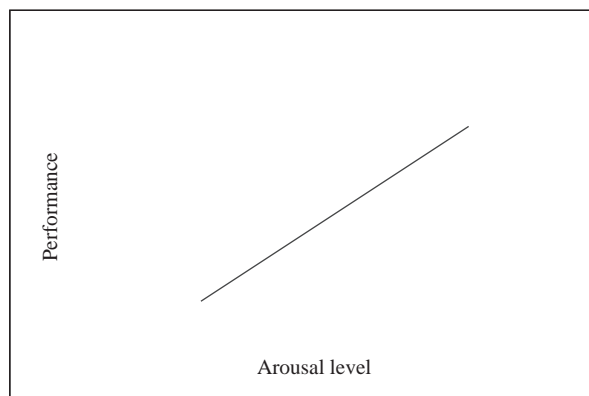
FIGURE 2   Drive theory.

well accepted for gross motor tasks requiring strength and speed, but it is less well accepted for tasks requiring balance, accuracy, and fine motor skills (e.g., putting a golf ball).

Drive theory is most beneficial to practitioners in understanding the link between skill level and optimal arousal. It is unlikely that an excellent athlete facing an easy task will garner sufficient emotional and physiological motivation to perform at his or her best. In such a situation, it benefits that athlete to set challenging process goals for the task at hand, in essence increasing the perceived challenge of the task. Conversely, regardless of the potential benefits of arousal to explosive movements and strength, a novice athlete will not reap great rewards from achieving exceptional arousal levels. Such an athlete should be counseled to find a state of calm that allows him or her to feel comfortable on the playing field or court.

4.2. Inverted-U Hypothesis

The inverted-U hypothesis states that moderate levels of arousal are ideal for optimal athletic performance

(Fig. 3). Performance gradually improves as a competitor goes from underarousal to the alert state of moderate arousal, after which point performance declines as the individual becomes overaroused.

Unlike drive theory, the inverted-U hypothesis does state that there is a point of arousal that exceeds what is needed for excellence on the playing field. Yet since the original model posed in 1908 by Yerkes and Dodson, it has been suggested that one must account for task difficulty, unique sporting demands, ability level of athletes, and individual differences when employing the inverted-U hypothesis. For example, it has been suggested that the peak of a golfer's inverted U ought to be at relatively lower levels of arousal, whereas a sprinter's inverted U will exist at high levels of arousal. In both instances, too little arousal or too much arousal is debilitative, yet the midranges of arousal differ due to sporting demands. Therefore, it is suggested that although the arousal–performance relationship is that of an inverted U, the placement of this inverted U on the arousal continuum ought to vary due to sport demands and individual differences.

The inverted-U hypothesis is perhaps most readily embraced by those in sport psychology practice. This is due in part to the fact that its simplicity (a unidimensional model as opposed to a theory that considers arousal and anxiety or multiple forms of anxiety) makes it easily understood by athletes and coaches. An athlete's ability to move into and out of ideal performance states is clear by drawing an upside down U on a piece of paper. This theory is also appreciated by practitioners because it recognizes the possibility of overarousal. This further illuminates the challenges posed by a "one size fits all" fire and brimstone pep talk.
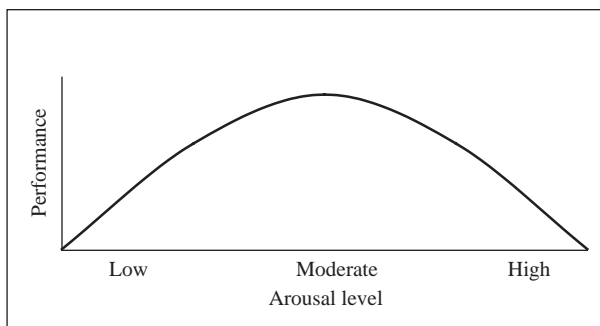
FIGURE 3 Inverted-U hypothesis.

4.3. Reversal Theory

Unlike other theories, the foundation of reversal theory is an athlete's interpretation of arousal rather than how high or low the energizing state is. Individual differences in perception determine how facilitative or debilitative arousal levels are. Depending on the situation that the athlete faces, arousal levels might be interpreted as "anxiety" or "excitement." For example, high levels of arousal might be interpreted by one athlete as lack of preparedness, whereas the same elevated arousal could be perceived as the appropriate energy for athletic success by another. Furthermore, at the root of this theory is the ability to change (i.e., reverse) one's interpretation of arousal. The reversal of interpretation might lie in improved skill levels (through practice), realization that tasks are not as challenging as thought previously, or support of coaches or teammates. Ultimately, according to this model, optimal performance is achieved when a player's preferred arousal level is in harmony with his or her actual arousal levels.

Sport psychology coaches who understand the reversal theory often see the importance of cognitive restructuring. The fact that perceptions of threats and opportunities on the playing field can change the facilitative nature of arousal makes an athlete's beliefs and thought processes essential for success. Regardless of the situation, an athlete nearly always has an opportunity to find his or her ideal arousal level. For example, a loud and hostile crowd could be perceived as an "extra player" for the home team, reducing the visiting team members' good feelings and confidence (i.e., debilitative perceptions). Conversely, the visiting team members could see and hear the hostile crowd and think, "We must be good! This crowd is worried and needs to help their team" (i.e., a facilitative perception). Drawing from the teachings of cognitive and/or rational–emotive therapies can be very useful to a practitioner who appreciates the reversible nature of arousal's effects on athletic performance.

4.4. Newer Theories

Other theories related to arousal in sport performance are the Individual Zones of Optimal Functioning (IZOF), the multidimensional anxiety theory, and the Cusp Catastrophe Model. Hanin suggested that in the IZOF, each individual has a certain level of emotion where he or she competes optimally. This "zone" of emotion can be high, low, or somewhere in between, depending on each person's unique

cognitive, physiological, and technical resources. The multidimensional anxiety theory suggests that cognitive anxiety (i.e., thoughts/worries) has an inverse linear relationship with performance, whereas somatic anxiety (i.e., physiological) displays an inverted-U relationship with performance. Lastly, the Cusp Catastrophe Model states that one must consider cognitive anxiety, physiological arousal, and performance demands simultaneously. In this theory, the relationship between anxiety/arousal and performance is not a neat linear one but rather one that can be filled with great peaks and valleys. In particular, one can achieve the greatest performances under high cognitive stress and arousal but, at the same time, can experience the greatest performance decline when things go awry.

5. EFFECTS OF AROUSAL ON ATHLETIC PERFORMANCE

Arousal levels directly influence physiological, cognitive, and motor modalities necessary for excellence on the playing field (Table II). Understanding these modalities

can help in the identification and management of under-arousal and overarousal. Physiologically, modalities measurable through psychophysiological monitoring, such as brain waves, skin conductance, heart rate, blood pressure, and muscle tension, all are elevated when an individual is aroused. Although these modalities are directly related to somatic arousal, each must be carefully measured because most are automatically increased during physical exertion. Perhaps the speed of recovery to "normal" levels between points and after play is when the goodness of one's arousal level is most apparent.

Arousal level also influences one's cognitive functioning. It is suggested that an athlete beyond his or her optimal arousal level is likely to suffer from poor decision making, impaired memory, and rushed thinking. Nideffer posited that in sport, heightened arousal levels also impair an athlete's ability to focus on the task at hand. Contrary to the lay assumption that over-aroused athletes think too much, it is suggested that being outside one's optimal arousal zone leads to thinking about the wrong things at the wrong time. In particular, Nideffer suggested that one begins to concentrate on thoughts, emotions, and feelings (specifically fear and worry). There is little to suggest that underarousal significantly influences one's cognitive

TABLE II
Effects of Arousal on Athletic Performance

Modality	Low arousal	High arousal
Brain wave (electroencephalography)	Increased alpha waves or theta	Increased beta waves
	Good performance on fine motor skills	Good performance on gross motor skills
Skin conductance	Lack of perspiration	Increased perspiration, increased conductance
	Good for a sport like golf	Good for high-energy sports
Heart rate	Decreased heart rate	Increased heart rate
	Ability to compete over long periods of time	Ability to perform skills with speed and power
Blood pressure	Decreased blood pressure	Increased blood pressure
	Effects uncertain	Effects uncertain
Muscle tension (electromyography)	Muscles relaxed	Increased muscle tension
	Fluid technique, potential poor form	Good strength and stability, potential for rigid movements
Focus	May be inattentive	Attention narrows, ultimately leading to focus on fears and worries
Memory	No deficit	Difficulty in remembering large chunks of information and complex tasks
Anaerobic power	Impaired strength and speed	Increased strength and speed
Fine motor control	Improved dexterity	Impaired

abilities, yet at low arousal levels few athletes are able to garner the necessary energy and alertness to succeed at the physical tasks of sport.

6. MANAGING AROUSAL LEVELS

There are many psychological techniques that allow one to effectively manage arousal levels. Understanding that arousal, anxiety, concentration, confidence, and physiological modalities all are inherently linked, it is clear that there are many approaches to help athletes find and maintain facilitative levels of arousal. Interventions can be broken down into three categories: cognitive, behavioral, and cognitive–behavioral. (Although dividing interventions into these categories is useful in understanding the practice of sport psychology, it is important to recognize that the mind and body are inherently linked. This reminds one that these categories are not mutually exclusive.)

6.1. Cognitive Interventions

Increasing an athlete's awareness about his or her optimal arousal levels is the foundation of an intervention. Developing awareness might be achieved through successive postperformance reflections that examine levels of excitement, concentration, emotions, and confidence during a competition. Similarly, a retrospective comparison of "best performances" versus "worst performances" can paint a powerful picture of an athlete's optimal arousal state.

Setting realistic and motivating goals gives an athlete the ability to energize while maintaining appropriate focus. It is appropriate to have a "mission" or goal for one's career and season as well as for each practice and competition. Goals direct an athlete (i.e., energizing) and clarify purpose (i.e., minimizing stress and relaxing).

Cognitive restructuring and self-talk can appropriately energize an athlete and calm him or her down when necessary. Cognitive restructuring fits nicely with the reversal theory of arousal, suggesting that modifying one's perspective and thoughts will help the athlete to find appropriate activation levels. Self-talk or the use of cue words can energize (i.e., "quick feet") or calm down (i.e., "slow and smooth") an athlete when necessary. Thought stoppage allows an athlete to catch himself or herself when he or she begins to focus on task-irrelevant cues.

Mental imagery, another cognitive intervention, is popular in the field of sport psychology. A good imagery session in which the body is relaxed and all of the senses are incorporated can be used to relax or energize an individual. Accessing one's memory bank of sporting highlights can serve as very motivational imagery, hence increasing arousal levels. Conversely, imagery can be used to escape from the stresses of competition, consequently reducing arousal.

6.2. Behavioral Interventions

There are a variety of relaxation techniques that can benefit an athlete who is overaroused. The following are a few techniques that are used regularly by sport psychology practitioners: controlled breathing, progressive muscle relaxation, meditation, relaxation response, and biofeedback. A few of these interventions can also be used by an athlete to energize, specifically quick shallow breathing and gaining understanding of excited physiological levels through biofeedback. Actions such as quick sprints, jumping rope, and dynamic stretching also serve to energize the underaroused athlete.

6.3. Cognitive–Behavioral Interventions

Both cognitive and behavioral techniques are combined during preperformance preparation and competitive routines. Prior to a competition, successful athletes commonly engage in activities such as setting goals, listening to music, using mental imagery, and physically warming up so as to have optimal arousal at the start of the competition. Similarly, while in the midst of play, athletes often use competitive routines to relax and refocus. Using a physical cue to put a prior play in the past, engaging in mental planning, and taking a good breath or two to help focus attention are often elements of such a routine. Routines are commonly used in sports where the flow of play is intermittent such as tennis, baseball, and golf.

See Also the Following Articles

Anxiety and Optimal Athletic Performance ■ Stress

Further Reading

Dienstbier, R. A. (1989). Arousal and physiological toughness: Implications for mental and physical health. *Psychological Review*, 96, 84–100.

Hanin, Y. (2000). *Emotions in sport.* Champaign, IL: Human Kinetics.

Kerr, J. H. (1985). The experience of arousal: A new basis for studying arousal effects in sport. *Journal of Sports Sciences*, *3*, 169–179.

Nideffer, R. M. (1989). *Attention control training for sport*. Los Gatos, CA: Enhanced Performance Services.

Spence, J. T., & Spence, K. W. (1956). The motivational components of manifest anxiety: Drive and drive stimuli. In C. D. Spielberger (Ed.), *Anxiety and behavior* (pp. 291–326). New York: Academic Press.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology of Psychology*, *18*, 459–482.

Zaichkowsky, L., & Baltzell, A. (2001). Arousal and performance. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (pp. 319–339). New York: John Wiley.

Assessment and Evaluation, Overview

Rocío Fernández-Ballesteros

Autonoma University of Madrid, Madrid, Spain

GLOSSARY

assessment contexts The fields in which psychological assessment is applied.

assessment process Series of tasks ordered in a specific sequence conducted by the assessor for solving the client/subject demand.

assessment targets Specific psychological characteristics (intelligence, cognitive skills, personality dimensions, behavioral responses, defense mechanisms, environmental conditions, etc.).

interview The interaction between two people in which one person (the interviewer) asks for and collects information with a given purpose or demand, and the other person (the interviewee, client, or subject) responds or answers questions.

observational method A set of procedures for collecting data, including a protocol with a pre-defined set of behavioral categories in which the subject's behavior is coded or registered.

projective methods Instruments containing relatively structure-free stimuli, considered especially sensitive to covert or unconscious aspects of behavior, that encourage a wide variety of idiographic subject responses with a minimum of subject awareness concerning the purpose of the test.

psychological assessment The discipline of scientific psychology devoted to the study of a given human subject (or group of subjects in a scientific applied field—clinical, educational, work, etc.), by means of scientific tools (tests and other measurement instruments), with the purpose of answering clients' demands that require scientific operations such as describing, diagnosing, predicting, explaining or changing the behavior of that subject.

psychological testing The process of administering, scoring, and evaluating tests or standard methods.

psychometrics The principles, theories, and techniques of test construction.

self-report A method for collecting data whose source is the subject's verbal message about him- or herself.

standard methods or tests Tests that have a standard set of stimuli, standard administration procedures, and a standard form of measuring subjects' responses, which can be transformed into standard scores.

1. INTRODUCTION

1.1. Concept

Psychological assessment is the discipline of scientific psychology devoted to the study of a given human subject (or group of subjects) in a scientific applied field— clinical, educational, work, etc., by means of scientific tools (tests and other measurement instruments), with

the purpose of answering client's demands that require scientific operations such as describing, diagnosing, predicting, explaining or changing the behavior of that subject. —Fernández-Ballesteros *et al.*, 2001.

Throughout the history of applied psychology, psychological assessment has been a key discipline, often to the extent of being confused with psychology itself and of constituting the public face of psychology. As several empirical studies have pointed out, assessment tasks are an ever-present element in all psychological work, from the scientific laboratory to applied settings. Thus, all types of psychological work involve assessment at some stage and to some extent, whether it be through the use of sophisticated data collection equipment, structured interviews, or qualitative observations, or by means of complex measures obtained through the administration of tests.

1.2. Brief History

Psychological assessment in some form or another can be found throughout the history of human thinking, in the Bible, in ancient China, and even in the works of Socrates, Plato, and Aristotle or those of the ancient physicians Hippocrates and Galen. Nevertheless, it is not until the late 19th century that we find scientific events related to psychological assessment. Table I offers an overview of the most important contributions to psychological assessment from the starting point of scientific psychology—the founding of Wundt's Psychology Laboratory in Leipzig in 1878—until 1950.

As can be seen in Table I, psychological assessment is a sub-discipline developed at the same time as experimental psychology. While psychology is devoted to the study of the principles of mind, consciousness, and behavior, psychological assessment is devoted to the study of how, or to what extent, these principles are fulfilled in a specific subject or group of subjects. For the study of a particular subject (or group of subjects), psychological assessment requires methods for data collection. Thus, psychological assessment is closely related to psychometrics and to psychological testing. Methodological advances in psychometric theory are important because all measurement devices are developed and evaluated using psychometric criteria; psychological testing and psychological assessment have even been confused. But psychological assessment cannot be reduced to either psychometrics or the processes of administering tests. Rather, as referred to in the above definition, psychological assessment always

involves a process of problem-solving and decision-making in which measurement devices and other procedures for collecting data are administered for testing hypotheses about a subject (or group of subjects), with the specific purpose of answering relevant questions asked by the client or the subject. The client and subject may be the same person or unit, or the client may be another professional or relative (e.g., a psychiatrist, an educator or a judge) who asks for the assessment of a third person or group of persons with a given purpose (description, classification, diagnosis, prediction, prognosis, selection, change, treatment, or evaluation).

1.3. Objectives and Contexts

But these applied questions require scientific objectives, of which the following are the most common: description, classification, prediction, explanation, control, and evaluation. By way of example: when a psychiatrist asks a psychologist for a diagnosis of a given patient, the psychologist should proceed to the description and classification of the patient; when a client asks a psychologist who is the best candidate for a job, the psychologist must predict the candidate who is likely to perform best; when a couple or a family asks for counseling, the psychologist should analyze the situation before intervening; when a pupil has a learning problem, the psychologist must analyze the case and hypothesize about the circumstances that explain or cause the problem, manipulate its cause(s), and evaluate the extent to which it has been solved.

As can be seen from the examples, these various objectives emerge from different applied settings, the three most common assessment contexts being clinical psychology, educational psychology, and work and industrial psychology.

Finally, the relevant events listed in Table I emerge from different models of scientific psychology, are related to specific psychological content or targets (intelligence, aptitudes, personality characteristics, etc.), and imply a variety of methods (interview, self-reports, projective techniques, etc.).

2. THEORETICAL PERSPECTIVES

Like psychology in general, psychological assessment developed in association with different models or theoretical frameworks, which have given rise to different theoretical formulations, targets, methods, and objectives. Table II shows the psychological models that have had

Principal Contributions in the History of Psychological Assessment from 1880 to 1950

Year	Author(s)	Contribution
1884	Francis Galton (1822–1911)	At London's South Kensington Museum, British scientist Galton set up the "Anthropometric Laboratory" for the assessment of physical and psychological characteristics (reaction time, visual acuity, discriminative speed and other sensory, perceptive, and motor measures).
1890	McKeen Cattell (1860–1944)	Working on his doctoral dissertation at Wundt's laboratory in Leipzig, the American Cattell developed the first definition of "mental test" as the measurement tool for higher mental processes.
1894	Emil Kraepelin (1856–1926)	One of Wundt's first pupils, Kraepelin proposed a comprehensive system for classifying psychopathology, taking into consideration mental states, memory, etc.
1896	Hermann Ebbinghaus (1850–1909)	Ebbinghaus defined intelligence as the ability to interrelate and combine correctly. He developed the first group intelligence test, the Completion test, consisting of a series of texts with words and parts of words omitted. Subjects (in a group) had to fill in as many blanks as possible in 5 minutes.
1896	Karl Pearson (1857–1936)	The Briton Pearson developed the mathematical method of correlation, a fundamental tool for the development of tests.
1898	E. L. Thorndike (1874–1949)	Thorndike reported the use of mazes and puzzles to measure the intelligence of cats. In 1904, he wrote *An Introduction to the Theory of Mental and Social Measurements*.
1904	Charles Spearman (1863–1945)	Two important developments can be attributed to this British scientist (who worked for several years in Germany): the development of reliability (*The Proof and Measurement of Association between Two Things*), and a theory of intelligence that led to the development of methods for locating a general factor underlying a group of tests.
1905	Alfred Binet (1857–1911)	Working on intelligence in an attempt to select students for compulsory education in France, Binet and his colleague Simon developed a set of situations that constituted the first test of intelligence.
1906	Karl Jung (1875–1961)	Jung introduced the first word association method for diagnosing psychopathology, using a list of 100 words, to each of which the subject was asted to respond as quickly as possible with the first word that came to mind.
1914	Robert S. Woodworth (1869–1962)	The first personality questionnaire was developed by Woodworth, intended as a screening process to identify those unfit for active service in World War I. The personal data sheet consisted of a set of self-reported items selected from psychiatric interviews.
1917	Arthur S. Otis (1886–1964) and Robert M. Yerkes (1876–1956)	Developed the "Army Alpha and Beta," general intelligence tests for the assessment and selection of American recruits for service in World War I.
1920	Kurt Goldstein (1878–1965)	Scientific neuropsychological assessment emerged in Germany after World War I from research on the concomitants of brain damage. After concluding that individuals with critical injuries display loss of ability for abstract thinking and that their behavior involves less generalization, Goldstein (in collaboration with Scheerer) developed several tests of abstract and concrete thinking.

Continued

Year	Author(s)	Contribution
1921	Hermann Rorschach (1884–1922)	The first projective assessment methods were developed by the Swiss Rorschach and published in his *Psychodiagnostik.* The method used 10 inkblots, to which subjects have to respond by saying what they "could be, or looked like." This method was designed to assess personality and psychopathology.
1934	Lev Vygotsky (1896–1934)	From the perspective of "Soviet psychology," the Russian Vygostky, in his *Thinking and Language,* developed the concept of the assessment of proximal development, which constitutes the roots of cognitive strategies of assessment.
1938	Henry A. Murray (1893–1988)	In the book *Exploration in Personality,* Murray (with Morgan) presented the thematic apperception test. Consisting of a set of ambiguous pictures with human figures, this test requires subjects to make up a story. The main objective of this method is to assess personality.
1943	Office of Strategic Services (United States)	Using methods first employed in Germany, but later developed in Britain, an important program for the global assessment of individuals, "Assessment of Man," was consolidated. As Dubois (1970) stated, "Instead of attempting to measure rather limited aspects of behavior as had been the case with the most psychometric and clinical devices up to that time, a program was designed to describe the way the individual was able to act in a wide variety of situations" (p. 110).
1950	Monty B. Shapiro (1951–)	At a meeting of the British Society of Psychology, Shapiro presented an important work developed in the Maudsley Hospital, "Experimental Approach to Diagnosis and Psychological Testing," which would serve as the basis for behavioral assessment and other idiographic developments.

most repercussion in psychological assessment: trait, medical, dynamic, behavioral, cognitive, and constructivist.

2.1. Trait

As human beings display wide variability in the way they think, feel, and behave, one of the objects of inquiry in psychology has been individual differences in human characteristics related to intelligence or mental aptitudes, affect and emotions (aggressiveness, hostility, etc.), personality (extraversion, flexibility, etc.), psychopathology (depression, anxiety, etc.). All of these have constituted important targets for assessment.

The trait approach uses several data collection techniques, from questionnaires or standardized tasks to rating scales. It is, therefore, the trait approach that has inspired the most well known psychometric tests. The main objective of this approach is to find the true score of a given subject in a specific trait or his or her true position in relation to a normative or standard group. In practical terms, these comparisons allow psychologists to describe and classify subjects with regard to a set of

traits or medical conditions and to predict their future behaviors. The explanation objective would not be applicable to the trait model, for reasons of tautology.

2.2. Medical

Historically, the development of psychological assessment has been closely related to clinical settings. Therefore, from the outset a medical model emerged, focusing principally on mental disorders. Several classification systems have been developed, the most well known being the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (currently the *DSM IV-R*) and the World Health Organization's *Classification of Mental and Behavioral Disorders* (currently the *ICD-10*), both of which constitute descriptive category systems with no etiological bases (neuropsychological illness and assessment are not included in this medical approach).

The purpose of these classification systems is to provide clear descriptions of diagnostic categories of mental disorders in order to aid clinicians in diagnosing

Theoretical approach	*Targets*	*Methods*	*Goals*
Trait	Trait	Questionnaires	Description
	Factor	Intelligence tests	Classification
	Dimension	Rating scales	Prediction, prognosis,
	Characteristic		or selection
	Syndrome		
Medical	Mental and behavioral disorders	Classification systems	Description
		Psychopathological tests	Classification and
			diagnosis
Psychodynamic	Unconscious mental process	Projective techniques	Description
	Conflicts	Completion methods (words, phrases,	Classification
	Defense mechanisms	stories, etc.)	Diagnosis
			Prediction
			Explanation
Behavioral	Behaviors:	Observation	Description
	Motor	Self-observations	Prediction
	Physiological	Self-reports	Functional
	Cognitive	Physiological measures	explanation
	Environment:		
	Physical		
	Social		
Cognitive	Cognitive processes	Performance on cognitive tasks and	Description
	Cognitive strategies	cognitive equipment	Prediction
	Mental representations	Thinking aloud	Explanation
	Other cognitive constructs	Self-reports	Evaluation
	Emotions verbally expressed		
Constructivism	Constructed knowledge	Narrative assessment	Description
	Personal constructs, personal	Autobiography	
	plans, meanings, etc.	REP-Grid	
		Personal documents	
		Content analysis	

and referring to these disorders. From this perspective, and with the aim of establishing objective procedures for assessing mental and behavioral disorders, several tests have been developed (e.g., the Multiphasic Minnesota Personality Inventory). In sum, description and classification are the two main objectives of the medical model.

2.3. Psychodynamic

Included under the title of psychodynamic are a series of relevant targets: unconscious mental processes, conflicting attitudes generating anxiety, defensive mechanisms, and interpersonal patterns governed by developmental experience. As stated by Weiner (2003),

the psychodynamic approach serves the purpose of elucidating "aspects of personality structure and personality dynamics in way that clarify the role of drives, conflict and defense, and object representations in shaping people are likely to think, feel and act" (p. 1012).

Throughout the history of psychological assessment, several instruments have been developed from a psychodynamic perspective. Perhaps the most well known are projective techniques using semi-structured visual materials (such as the Rorschach tests or the thematic apperception test), drawing a specific object (draw-a-person, draw-a-tree, draw-a-family), and finally, approaches using verbal material (words, sentences, stories, etc.). The goals of the dynamic model

include description and diagnosis as well as prediction and explanation.

2.4. Behavioral

The behavioral assessment approach has been defined in several ways throughout the 20th century, but three fundamental pillars of these definitions can be identified and reduced to the following: (1) human behavior can be explained by the interaction between the person (from a biological and psychological point of view) and the environment; (2) there are three different modes or systems: motor, cognitive, and psychophysiological; and (3) psychological problems have multiple causes, essentially environmental. Thus, as Table II shows, targets in behavioral assessment are motor, cognitive, and physiological responses, in addition to environmental conditions.

A methodological condition of behavioral assessment is that all targets should be assessed through the multimethod approach. Thus, assessment of behavioral and environmental conditions and their respective interactions takes place through observation (direct observation as well as observations by relatives), self-observation, self-report, and psychophysiological techniques. Finally, because behavioral assessment emerges from a clinical perspective, its objectives are description, prediction, functional explanation (need for treatment), and treatment evaluation. The main goals of the behavioral psychologist are related to the environmental or personal conditions that can be functionally linked to target behaviors and provide indications for the planning of treatment.

2.5. Cognitive

The cognitive model refers to two theoretical perspectives, the first related to the study of cognitive processes and the second associated mainly with therapy and treatment, under the assumption that cognitive processes or structures (or verbal expression of emotions) are causal aspects to be taken into account if the intention is to change behavior.

Table II shows the most important targets for the cognitive model: cognitive processes (such as attention, perception, learning and memory, language, and imagery), mental representations (e.g., visuospatial representations, automatic thinking), cognitive structures (cognitive abilities), and other cognitive constructs or forms of emotional expression. These processes, strategies or structures are assessed mainly through subjects' performance in cognitive tasks. For example, attention

can be measured by asking the subject to find a specific letter (e.g., "m") in a text, or to press a button when a specific cue appears on a screen (testing reaction time).

Cognitive processes or strategies can also be measured using meta-components. Thus, the subject may be asked to "think aloud" while solving a given task. The cognitive model also considers the meta-component of emotions, so that subjects may be asked to self-report what they feel in a particular situation. Finally, the objectives of the cognitive model—in the same line as those of the behavioral model—are the following: description (how the subject can be described on the basis of these processes and structures), prediction (what the subject can be expected to do on the basis of these processes), explanation (which processes may be controlling targets, so that they can be manipulated in order to change targets), and evaluation (to what extent treatment has produced the expected changes).

2.6. Constructivist

The constructivist theoretical approach to psychology stresses conceptual, epistemological, methodological, and ontological characteristics, as opposed to an objective and realistic view of the science. From this approach, all knowledge is considered as a social product, so that there cannot be any universal law or principles: subjective aspects and local custom or practices dictate. Therefore, targets must be selected for each individual and his or her specific way of constructing social reality, the meaning of life, or personal plans, or of producing personal constructs. Methods are idiographic rather than standard. Finally, utility is the most important characteristic of any assessment, as opposed to validity. Table II shows these idiographic targets making up the basic personal knowledge developed across the subject's life span.

Several methods have been developed within the constructivist approach, all of which attempt to capture the essence of the person in question. Perhaps the most well known is the repertoire grid technique. Originally developed by Kelly in the 1950s, several different versions have appeared over the last half-century, all with the basic purpose of assessing the subject's personal constructs. Other methods involve the use of narratives developed by different authors, through which a sense of identity, coherence, meaning of life, and other personal phenomena can be assessed. With these narrative-based methods, the use of autobiography is an essential element for assessing these constructs. Finally, the data from these methods is analyzed through content analysis, which is

essential to constructivist methodology, because herme-neutics is, basically, a system for discovering targets.

3. METHODS AND TECHNIQUES

Marx and Hillix made a distinction between scientific method, methods, and techniques (methodic tech-niques or instruments). While the scientific method is the basic strategy for the advancement of science (e.g., the hypothetico-deductive method), methods are gen-eral data collection procedures, and techniques can be defined as those specific instruments or tests developed for collecting data in a particular field.

An important body of knowledge within psychologi-cal assessment is the set of methodic techniques for data collection. However, before going on to describe these methods, it is necessary to first identify some general conditions of psychological assessment methods.

1. Methods, techniques, and instruments should not be confused with psychological assessment. As stated previously, methods and tests constitute the tools for data collection and testing hypotheses, while psycho-logical assessment is a lengthy process of problem-solving and decision-making.

2. There are standard and non-standard methods. All techniques and instruments have a set of stimuli (ques-tions, tasks, etc.), a set of administration procedures, and a set of responses; finally, subjects' responses can be transformed into scores. Standard methods or tests are those that have a standard set of stimuli, standard administration procedures, and standard forms of obtaining subjects' responses—responses that can be transformed into standard scores. Standard methods usually have sound scientific properties and present psychometric guarantees. These standard methods, which allow between-subject comparisons, are known as psychometric tests. But not all data collected in psy-chological assessment derive from standard methods.

3. As described above, assessment targets are spe-cific psychological characteristics (intelligence, cogni-tive skills, personality dimensions, behavioral responses, defense mechanisms, environmental conditions, etc.) with several degrees of inference. That is, the majority of these targets are psychological concepts that are operationalized by means of a given method, tech-nique, or measurement device. Thus, all methods in psychology can be considered procedures for operation-alizing a given psychological construct or target for assessment. Because each method has its own source

of error, the best way to conduct a rigorous assessment is to use several methodic techniques, tests, or mea-sures or, in other words, to triangulate a given concept through multimethod. For example, if in assessing a particular personality construct (e.g., extroversion) the subject can be administered a standardized self-report, he or she can be given a list of adjectives, and his or her behavior can be observed in an analogue situation or by administering a projective technique.

With these three general conditions in mind, consider the broad categories in which psychological assessment methods can be classified: interview, observation, self-reports, and projective techniques. It should be borne in mind, however, that these four categories can them-selves be divided into more specific sub-categories.

3.1. Interview

Interviewing is the most commonly used method in psychological assessment in all basic and applied fields: clinical, educational, work and organizations, forensic, laboratory, and so on. It has the broadest scope, so it can be used for assessing any psychological event (subjective or objective, motor, cognitive or physiological).

For assessment purposes, an interview is basically an interactional process between two people in which one person (the interviewer) elicits and collects informa-tion with a given purpose or demand, and the other person (the interviewee, client, or subject) responds or answers questions. The information collected can refer to the interviewee (motor, cognitive, or emotional responses) or to relevant others. Thus, the interview can be thought of as a self-report or as a report to or by another.

Because the interview is an interpersonal process, it has the following several phases:

1. Interview preparation: Before starting an inter-view, the psychological assessor should plan the inter-view. What are the purposes of the interview? What are the subject's characteristics? What format will the in-terview take? How long will the interview last?

2. Interview starting point: A brief presentation by the interviewer about the purpose of the interview, what the interviewer is expecting, and the plans for the assessment process.

3. Body of the interview: The interview is always led not only by the purpose of the assessment but also by the theoretical approach of the interviewer. These two prin-cipal conditions influence the content (what type of

question) and the format (amount of structure in questions and answers) of the interview. By way of example, the main common and shared areas of an interview are as follows: sociodemographics, interview purpose or demand, family and social condition, education, work, interest and hobbies, health, beliefs, and values.

4. Closing the interview: The interviewer closes the interview by briefly summarizing the data collected and describing the subsequent steps or actions.

There are several typologies of the interview depending on: (1) the structure of the questions and answers (structured, semi-structured, non-structured), (2) the context in which the interview takes place (clinical, personnel selection, counseling, forensic, etc.), (3) interviewer style (directive, non-directive), (4) theoretical perspective (dynamic or psychoanalytic, behavioral, phenomenological, etc.), and (5) number of subjects assessed (individual, group, family, team group, etc.).

The interview as an assessment method has several advantages and limitations, such as the following:

1. It offers a situation in which subjects can act and interact with several verbal stimuli (or questions) and can be observed. Also, in this situation several interpersonal emotional relationships (interviewer–interviewee) develop that are highly relevant to forthcoming psychological actions.

2. The interview is a broad technique for collecting data; that is, any condition or purpose (past or present; internal or external; descriptive, classificatory, or explanatory, etc.) can be assessed. Also, this flexibility of the interview allows psychologists to clarify a given situation or action of the subject.

3. However, information collected in interviews may present considerable bias and error deriving from the interviewee, the interviewer, and the situation. Inter-interviewer agreement and test-retest reliability is relatively low, and validity depends on the construct assessed.

4. Finally, as emphasized by Meyer et al., although most types of psychological testing—including the interview—show evidence of their predictive validity, the use of multimethods is essential in order to gain a complete understanding of each case.

3.2. Observational Methods

All psychological assessment contains observations. Indeed, as Anastasi (1988) stated: "a psychological test is essentially an objective and standardized measure of a sample of behavior" (p. 23). Thus,

observation is one of the broadest categories in psychological assessment. Common intelligence tests (such as the Wechsler Intelligence scales), as well as cognitive or psychophysiological measurement using special equipment, can be used as observational procedures. In brief, an observational method consists of a set of data collection procedures containing a protocol with a predefined set of behavioral categories in which subjects' behaviors are coded or registered. These behaviors are external, motor events observed in natural, analogue, or test situations; observation may sometimes involve the use of sophisticated equipment such as tachistoscopes, polyreactigraphs, or psychophysiological instruments that allow response amplification.

Observational methods have been classified depending on the following: (1) the observer: participant versus non-participant observer; (2) the target or observed event: overt event (motor response or stimuli parameter) versus subject's attribute or molar characteristic of the observed situation; (3) the situation in which observations take place: natural versus artificial or standardized situations (or even with equipment amplifying the target); and (4) the protocol: coding system versus open description of the continuum of behavior. Table III shows examples of this classification system.

Observer, subjects and targets, and observational protocol are sources of bias of the observational assessment method. Below is a summary of how to maximize the accuracy of observational methods:

1. *Observer.* Observational methods in psychology should maximize objectivity through the assessment of overt (observable) events, ongoing observer training, use of several observers, and checking of inter-observer agreement.

2. *Subject observed.* One of the most important biases of observational methods is subject reactivity. Reactivity can be minimized in various ways, such as use of natural settings, subjects being unaware, and long habituation periods.

3. *Observational systems.* Coding or registration systems require well-defined targets and procedures, a small number of observation categories, and standard data about psychometric soundness.

3.3. Self-reports

As defined elsewhere, the self-report is a method for collecting data whose source is the subject's verbal message about him- or herself. Self-reports are the most

TABLE III
Type of Observational Methods and Examples of Observational Techniques

Observer	Target	Situation	Protocol	Example
Nonparticipant Expert observes target behaviors or attributes	Motor behaviors Product of behavior Attributes	Natural situation registered through video, film, two-way mirror, etc.	Continuum of observed behavior Category system Behavior checklist Rating scales	Standardized observational code (SOC-III) (Cerezo *et al.*, 1991)
Participant but expert	Motor behavior Product of behavior (dimensions) Verbal behaviors Attributes	Standardized/Non-standardized situations	Category system Behavior checklist (occurrence, frequency accuracy, reaction time, duration, etc.) Rating Scales (Likert-type scales)	WAIS III (Wechsler, 2000)
Participant and/or relatives	Dimensions of motor and verbal behavior Attributes	Natural situation Non-standardized	Behavior checklist Rating scales	Eyberg behavioral inventory (Eyberg & Ross, 1978)

common assessment procedures for collecting data in psychology and psychological assessment. Self-reports provide information about thousands of targets—from historical events to motor behaviors in the present situation, even subjects' future plans. In fact, human beings can report on their private, unobservable, subjective world, their thoughts, and their emotions, as well as public observable events. This means that, in answering a verbal question, subjects may carry out certain mental operations (from remembering what they did a few minutes previously to inferring why they acted in a certain way a long time ago). Such operations interfere with the fidelity of self-reports.

Also, when a self-report or a set of self-reports is collected, psychologists can either take them as they are (from an isomorphic perspective) or make inferences from them after carrying out statistical manipulations (e.g., item aggregation or item anchoring in new scales). In other words, in order to test theoretical hypotheses, psychologists interpret the self-reports as indices of internal structures or processes, such as attitudes, personality characteristics, motives, and other intrapsychic concepts developed from trait theory, medical models, cognitive psychology, and/or dynamic theories. Thus, since the very beginning of scientific psychology, aggregates of self-reports assessed via questionnaires through inter-subject designs (subject aggregates) have been

used as one of the main assessment methods for describing, classifying, predicting, and even (tautologically) explaining human behavior.

In psychological assessment, clients' self-reports are collected in different formats. As already mentioned, interviews frequently represent the starting point and conclusion of a professional relationship in clinical, educational, and work settings. When self-reports are used as formal procedures for data collection, they imply a broad methodological category: interviews always involve self-reports, and self-observation and self-monitoring also require self-reports. Finally, self-reports can be obtained through questionnaires, inventories, or scales containing a set of relevant verbal statements and a variety of response formats (true or false options, checklists, scaled responses, etc.).

As emphasized elsewhere, self-reports have several sources of bias. Whatever the type of event reported—from objective events to subjective ones—several internal and external conditions, such as length of the question, formulation of the question, type of answer and response requested, and subject's characteristics, strongly affect the fidelity of self-reports. Table IV shows examples of general issues related to self-reports.

In summary, human beings are able to report on a myriad of events and are an important source of information; several assessment methods (such as interviews

TABLE IV

Examples of General Issues in Self–Reports[a]

Targets	Subject mental operations	Psychologist inferences	Response format	Type of self-report
Motor response, external events	Accessibility to the target	Self-report as data vs	Open	Interview
Cognitive behavior	Level of inference of required transformation	Self-reports as sign	YES/NO, multiple choice, rating-scale	Questionnaire, inventory, scales
Subjective appraisal				
Psychophysiological responses	Time: Past, present, future		Latency, occurrence, frequency, amplitude, etc.	Self-observation, self-monitoring

[a]Modified from Fernández-Ballesteros & Márquez, 2003.

and self-observation, self-monitoring, questionnaires, inventories or scales) are based on self-reports. Even though self-reports have important sources of errors, so that self-report accuracy is threatened by several types of response distortions (positive and negative impression management, faking, acquiescence, and social desirability, among others), these can, at least partially, be measured and controlled. Finally, it is commonly accepted that self-report questionnaires can be optimized through various strategies, and that their accuracy can be maximized. When self-reports are taken as a source of behavioral information, the level of inference should be as low as possible, items should be as direct as possible, questions should be specific and with situational reference, the level of transformation of the information requested should be as low as possible, and finally, the information requested should refer, preferably, to the present.

3.4. Projective Techniques

According to Pervin (1975), a projective technique "is an instrument that is considered especially sensitive to covert or unconscious aspects of behavior, permits or encourages a wide variety of subject responses, is highly multidimensional, and evokes unusually rich or profuse response data with a minimum of subject awareness concerning the purpose of the test" (p. 33).

The basic assumptions of projective techniques are closely related to the psychodynamic model and psychoanalytic theory: (1) they involve the presentation of stimuli relatively free of structure or cultural meaning, consisting of inkblots, pictures, incomplete verbal sentences or stories, or performance tasks; (2) these

materials allow subjects to express idiosyncratic and holistic aspects of their personalities, and allow the assessment of dynamic constructs as well as defense mechanisms against anxiety, and conscious as well as unconscious processes and structures; and (3) subjects are not aware of the purpose of the test and the meaning of their responses to the test. Table V shows four main categories of projective techniques that can be identified depending on the material used and the task involved.

Although projective methods have been widely used, they have received considerable criticism, and continue to cause controversy. In fact, all the characteristics of projective methods—ill-defined stimuli, idiographic responses, non-standardized procedures, and subjective evaluation—have contributed to a lack of psychometric soundness. In sum, much more research is required for projective methods to be used as scientific tools. Nevertheless, there is an exception, as Weiner (2001) has emphasized: "the Rorschach Inkblot Method has been standardized, normed, made reliable, and validated in ways that exemplify sound scientific principles for developing an assessment instrument" (p. 423).

4. THE PROCESS OF ASSESSMENT, TREATMENT, AND EVALUATION

As mentioned previously, assessment should not be confused with its methods and tests, nor with a specific applied setting (such as the clinical one). It begins when a client makes a specific demand of a psychologist in a given context. This demand triggers a long process of problem-solving and decision-making that requires a series of tasks ordered in a specific sequence. This sequence

Types and Examples of Projective Techniques Depending on Materials Used and Task

Type of technique	Material	Task	Examples
Structural	Inkblots	Describe what it looks like	Rorschach Inkblots test (Rorschach, 1929)
Thematic	Pictures	Tell a story	Thematic apperception test (TAT) (Murray, 1936)
Expressive	White paper and pencil	Draw a picture	Machover Draw-a-Person test (Machover, 1948)
Associative	Words, incomplete sentences or stories	Complete the sentences or stories	Free association test (Rapaport, 1965)
Performance	Constructional elements (houses, trees, fences, etc.)	Build a town	Town test (Arthur, 1949)

is dictated by the objectives the psychologist must fulfill: description, classification/diagnoses, prediction/prognosis, explanation, and control. When the demand is for control, implying the need for an intervention, evaluation is an operation to be performed for appraising the treatment, or testing the extent to which the targets have changed in the expected direction. Thus, when assessment is conducted in order to produce change and is followed by treatment, evaluation should be the final step in the assessment process.

After 30 years of research into the assessment process, guidelines for the assessment process (GAP) have been set up in order to formalize it and published so that they can be discussed by the scientific community. The GAPs have been developed with the following purposes: (1) to assist psychological assessors in their efforts to optimize the quality of their work; (2) to assist the client coming from outside psychology in evaluating assessment tasks by allowing quality control; and (3) to facilitate the teaching of assessment, the standardization of practical considerations, and the design of advanced professional training programs.

Table VI presents the general framework of the assessment process. As it can be seen, the assessment, treatment, and evaluation process has four main steps: (1) analyzing the case; (2) organizing and reporting results; (3) planning the intervention; and (4) evaluation and follow-up. The two first phases deal with case formulation as well as with integrating, reporting, and discussing the case, and therefore refer to all assessment contexts (clinical, educational, work and industry, etc.), because description, classification/diagnosis, and prediction/prognosis are the assessment objectives in all cases. The third and fourth phases refer to those cases

demanding control and change and requiring explanatory assumptions, experimental designs, and, finally, treatment evaluation. Each step has several substeps, and from the substeps a set of 96 guidelines has been developed. The GAPs are procedural suggestions and should be taken as recommendations for helping assessors to cope with the complexities and demands of assessment processes in various applied contexts.

5. CONCLUSIONS

Psychological assessment is a subdiscipline of scientific psychology devoted to the study of a given subject with different purposes (describing, diagnosing, predicting, etc.) and in different settings (clinical, educational, etc.). Psychological assessment should not be confused with its methods or tests, since it involves a lengthy and complex process of decision-making and problem-solving that begins with the client's demands. There are several theoretical models proposing specific targets, methods, and assessment objectives.

Psychological assessment uses a variety of methods: interview, observation, self-reports, and projective techniques. Some of these have standard sets of stimuli, administration procedures, and ways of measuring subjects' responses and transforming them into standard scores; these methods are the well-known psychometric tests, but they are not the only means of collecting data in assessment. All assessment methods have particular sources of error, and all can be considered as fallible. The use of multimethod is highly recommended. Through the complex process of decision-making and problem-solving involved in

TABLE VI
Assessment Process: Basic Framework and Steps

START OF PROCESS: Process begins when a person/institution (client) asks an assessor to answer a question or give professional advice about a subject/single case.

PRECONDITIONS: Assessor examines whether he/she is qualified to satisfy the demand and whether the demand is in line with ethical criteria.

0. GENERAL PRINCIPLES

1. ANALYZING THE CASE (DESCRIPTIVE ASSESSMENT)

  1.1 ANALYZING DEMANDS, COMPLAINTS, AND/OR GOALS: Gathering general information about client's and/or subject's referral question.

    1.1.1 Investigating and evaluating client's and/or subject's questions.

    1.1.2 Synthesizing client's demands and aspects of the general problem situation.

    1.1.3 Formal agreement.

  1.2 FORMULATING TESTABLE ASSESSMENT HYPOTHESES ABOUT THE CASE: Converting demands, complaints, and/or goals into testable hypotheses.

    1.2.1 Formulating questions in technical terms, based on information gathered.

    1.2.2 Operationalizing technical terms by means of assessment procedures.

  1.3 COLLECTING INFORMATION: Gathering information relevant to the questions.

    1.3.1 Planning administration of assessment procedures.

    1.3.2 Applying assessment procedures.

    1.3.3 Evaluating the application of assessment procedures.

  1.4 INFORMATION PROCESSING, RELATING DATA COLLECTED TO THE ASSESSMENT QUESTIONS

    1.4.1 Analyzing data.

    1.4.2 Drawing assessment conclusions.

2. ORGANIZING AND REPORTING RESULTS: Technical preparation of results and reporting to client/subject.

  2.1 INTEGRATING RESULTS: Answering the client's/subject's questions as completely as possible.

    2.1.1 Combining results into a comprehensive case formulation.

    2.1.2 Formulating conclusions with respect to client's/subject's demands.

  2.2 REPORTING: Written and/or oral presentation of reports.

    2.2.1 Requirements for report generation.

    2.2.2 Including relevant information in the report

    2.2.3 Making report understandable.

  2.3 DISCUSSING AND DECIDING

    2.3.1 Discussing report with client, subject, and/or relevant others.

    2.3.2 Analyzing whether general circumstances warrant stop, re-start, or moving on to an intervention.

3. PLANNING THE INTERVENTION: If the assessor considers that an intervention is required, several assessment operations are necessary before treatment administration.

  3.1 SELECTING AND TESTING SPECIFIC INTERVENTION HYPOTHESES:

    3.1.1 Selecting and operationalizing intervention and outcomes variables.

    3.1.2 Reviewing and deciding on intervention procedures that best fit the single case.

    3.1.3 Selecting and assessing relevant variables for monitoring.

4. EVALUATION AND FOLLOW-UP: If an intervention has been carried out, several assessment operations should be conducted.

  4.1 COLLECTING DATA ABOUT EFFECTS OF INTERVENTION:

    4.1.1 Inspecting already-available data.

    4.1.2 Collecting post-intervention data.

  4.2 ANALYZING INTERVENTION OUTCOMES:

    4.2.1 Drawing conclusions from data gathered on the effects of the intervention.

    4.2.2 Reporting results to client, subject, and/or relevant others.

    4.2.3 If necessary, written report to client, subject, and/or relevant others.

Continues

Continued

4.3 FOLLOW-UP:

    4.3.1 Planning follow-up in agreement with client and/or subject.

    4.3.2 Assessing subject according to the established plan.

    4.3.3 Analyzing results.

    4.3.4 Discussing results with client, subject, and/or relevant others.

    4.3.5 If necessary, written report to client, subject, or relevant other.

END OF PROCESS: The assessment process comes to an end, if the assessor terminates the professional relationship with the person/institution (client) and subject/single case with respect to the assessment tasks.

---

psychological assessment, psychologists test hypotheses about the case, achieve assessment objectives, and draw certain conclusions in an attempt to respond to the demands of the client.

See Also the Following Articles

Clinical Assessment ■ Diagnostic and Statistical Manual of Mental Disorders ■ Intelligence Assessment ■ International Classification of Diseases (WHO) ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Psychophysiological Assessment ■ Traits

Further Reading

Arthur, H. (1949). *Le village. Test d' Activité Créative*. Paris: Pub.

Baer, R. A., Rinaldo, J. C., & Berry, D. T. R. (2003). Self-report distortions. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (Vol. 2, pp. 861–866). London: Sage.

Cook, T. D. (1985). Postpositivist critical multiplisms. In L. Shotland, & M. M. Marck (Eds.), *Social science and social policy*. Beverly Hills, CA: Sage.

Cerezo, M. C., & Whaler, R. G., *et al.* (1991). *SOC-III Standardized Observational Code*. Madrid: MEPSA.

Cronbach, J. L. (1975). Five decades of public controversy over mental testing. *American Psychologist*, *30*, 1–4.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis. Verbal reports as data*. Cambridge, MA: MIT Press.

Eyberg, S. M., & Ross, M. (1978). Assessment of child behavioral problems. *Journal of Clinical Psychology*, *3*, 113–116.

Fernández-Ballesteros, R. (2002). Behavioral assessment. In Smelser, N. J., & P.B. Baltes (Eds.), *International encyclopedia of social and behavioral sciences*. New York: Pergamon.

Fernández-Ballesteros, R. (2003). Self-report questionnaire. In S. N. Haynes & E. Heiby (Eds.), *Comprehensive handbook of psychological assessment* (Vol. 3, pp. 194–221). New York: John Wiley & Son.

Fernández-Ballesteros, R., De Bruyn, E. E. J., Godoy, A., Hornke, L., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H., & Zaccagnini, J. L. (2001). Guidelines for the Assessment Process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, *17*, 187–200.

Fernández-Ballesteros, R., & Márquez, M. O. (2003). Self-reports. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (pp. 871–877). London: Sage.

Freixas, G. (2003). Subjective methods. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment*. London: Sage.

Goldfried, M. R., Stricker, G., & Weiner, I. (1971). *Rorschach handbook of clinical and research application*. Englewood Cliffs, NJ: Prentice Hall.

Goldstein, G., & Hersen, M. (1990). Historical perspectives. In G. Goldstein, & M. Hersen (Eds.). *Handbook of psychological assessment* (pp. 3–21). New York: Pergamon Press.

Goldstein, G., & Beers, S. R. (Eds.). (2003). *Intellectual and neuropsychological assessment*. New York: Wiley.

Haynes, S. N., & O'Brien, W. H. (2000). *Principles and practice of behavioral assessment*. New York: Kluwer.

Haynes, S. N., & Heiby, E. M. (Eds.). (2003). *Behavioral assessment*. New York: Wiley.

Machover, K. (1948). *Personality projection in the drawing of the human figure*. New York: Ch. C. Thomas.

Marx, M. H., & Hillix, W. A. (1967). *Sistemas y teorías psicológicas contemporáneas. (Contemporary psychological theories and systems)*. Buenos Aires: Paidós.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment. *American Psychologist*, *56: 128–165*.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Murray, A. H. (1936). *Exploration in personality*. New York: Oxford Press.

Neimeyer, G. J. (Ed.). (1993). *Constructivist assessment*. Newbury Park: Sage.

Pervin, L. A. (1975). *Personality, assessment, and research*. New York: John Wiley & Sons.

Rapaport, D. (1965). *Test de diagnóstico psicológico* (Manual of Diagnostic testing). Buenos Aires: Paidós.

Rorschach, H. (1921). *Psychodiagnostik*. Berne: Hans Huber.

Segal, D. L., & Frederick, L. (2003). Objective assessment of personality and psychopathology: An overview. In Segal, D. L., & Hilsenroth, M. (Eds.), *Personality assessment*. New York: Wiley.

Sundberg, N. (1976). *Assessment of persons*. New York: Pergamon.

Schwartz, N., Park, D. C., Knauper, B., & Sudman, S. (1998). *Cognition, aging and self-reports*. Ann Arbor, MI: Psychology Press.

Scriven, M. (1991). *Evaluation thesaurus*. London: Sage.

Weiner, I. (2001). Advancing the science of psychological assessment: The Rorschach Inkblot method as exemplar. *Psychological Assessment*, *13*, 423–433.

Weiner, I. (2003). Theoretical perspective: Psychoanalytic. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment*. London: Sage.

Wiggins, J. (1973). *Personality and prediction. Principle of personality assessment*. Reading, MA: Addison-Wesley.

Assessment in Sport Psychology

Marc-Simon Sagal, Paul Thomas Sagal, and Geoffrey E. Miller

Winning Mind, LLC, San Diego, California, USA

GLOSSARY

applied sport psychology The application of psychology, typically toward the improvement of performance, gained through the study of the psychological factors that affect and are affected by participation and performance in sport.

consensual validation The process of confirming, through interview and behavioral observation, the existence of certain psychological qualities previously identified through formal assessment.

response set A situation-specific approach employed by an individual undergoing a psychological assessment to manipulate resulting assessment data (e.g., faking good).

response style An approach that reflects cross-situational, core personality characteristics and that can have an equally strong impact on assessment results (e.g., pessimism).

sport psychological assessment The systematic attempt to measure those psychological characteristics that influence performance and behavior in sport contexts.

sport psychology professional (SPP) An individual with appropriate academic credentials and experience who applies this expertise responsibly in sport settings; SPPs are interested in helping individuals to improve performance, appreciate the overall sporting experience, and enhance personal development and well-being.

sport-specific measures Inventories that attempt to elicit responses directly relevant to performance in sport generally or in one sport specifically.

trait A relatively enduring characteristic or habit.

state An occurrent characteristic (i.e., situationally dependent).

Assessment enables sport psychology professionals (SPPs) to obtain crucial performance-relevant information. The principal aim is to help athletes perform at an optimal level. To this end, both formal and informal methods are employed. These include interviews, behavioral observation, psychological testing, surveys, and inventories. Effective assessment requires sensitivity to test validity and reliability, response sets and styles, consensual validation, and ethical requirements. This article includes a brief history and a discussion of future directions.

1. INTRODUCTION

At its most basic level, assessment involves information gathering, interpretation, and evaluation. We all constantly engage in one form of assessment or another. It is how we make sense of the world. Sport psychology assessment is the systematic attempt to measure

psychological characteristics with the aim of improving sport experience and performance. Within the field of applied sport psychology, the method of assessment typically depends on the context (e.g., purpose, clients, time constraints). Although most sport psychology professionals (SPPs) associate assessment with more formal approaches (typically paper-and-pencil questionnaires and now online questionnaires), there are a variety of ways in which information can be collected and distilled. This article focuses on the practical applications of assessment and its use in applied settings. It also provides general information on selected assessment methods and a number of widely used instruments and tests.

2. ROLE OF ASSESSMENT

SPPs use assessment to gain crucial information about the people with whom they work. The aim is to help each of these individuals reach the level of performance desired, whether this is the level of a professional or Olympic athlete, an amateur competing in organized club competition, or a purely recreational sport enthusiast. But what exactly does an SPP need to understand about an individual to help? The answer to this question is controversial. Some SPPs believe that a clinical understanding of the client is necessary. Most SPPs focus on gaining more direct performance-relevant information (e.g., how the athlete pays attention, under what situations performance breaks down). Even if there were agreement on exactly what should be assessed, it would be difficult to get agreement on assessment methods. This difficulty increases where there is no agreement on which factors are most important.

There are four primary applications of applied sport assessment:

- Performance enhancement (individual and team)
- Athlete selection and screening
- Injury recovery
- Sport enjoyment

Some SPPs believe that all applications can be subsumed under the primary category of performance enhancement. Others emphasize the importance of having a positive sporting experience regardless of performance issues. Most people choose to work with SPPs because they want to be better at their sport. The presenting issue may be "recovering quickly from an injury," "resolving conflicts among team members," or "identifying mental strengths and liabilities," but the bottom-line concern is with performance (Table I).

TABLE I
Common Reasons for the Use of Assessment

1. To measure variables that are important to performance, sport enjoyment, and a variety of other factors that contribute to the quality of the sport experience. Examples include the following:
    a. To determine current mental skill ability
    b. To determine presence of factors associated with overtraining
    c. To monitor anxiety
    d. To assist in the presence of clinical issues
2. To reduce the amount of time it would ordinarily take to collect this relevant information through other means (e.g., observation)
3. To use a repeatable method for gaining critical information to help athletes
4. To quantify relevant information for analysis over time or between populations
5. To determine whether new factors or groups of factors are more or less important contributors to desired outcomes, states, or behaviors

3. METHODS OF ASSESSMENT

SPPs apply a variety of assessment methods to understand, predict, and modify behavior, including the following:

- Interviews
- Behavioral observation (including video review)
- Psychological testing, surveys, and inventories
- Third-party anecdotal data collection (from parents, coaches, teammates, etc.)

Typically, context will help to determine the most appropriate methods to employ. In most situations, SPPs will gather information using more than one of these methods.

3.1. Interviews

The most common form of assessment takes place through the process of formal and informal interviewing. There are typically three purposes for conducting an interview: to identify performance issues, to determine ways in which to intervene, and to validate inventory and questionnaire data. The interview itself can be structured by a set of preestablished questions and topics, or it can be free-flowing with a more open and dynamic feel. Topics and lines of inquiry depend on many factors, including the type of SPP–athlete relationship, the phase of the intervention/program, the severity
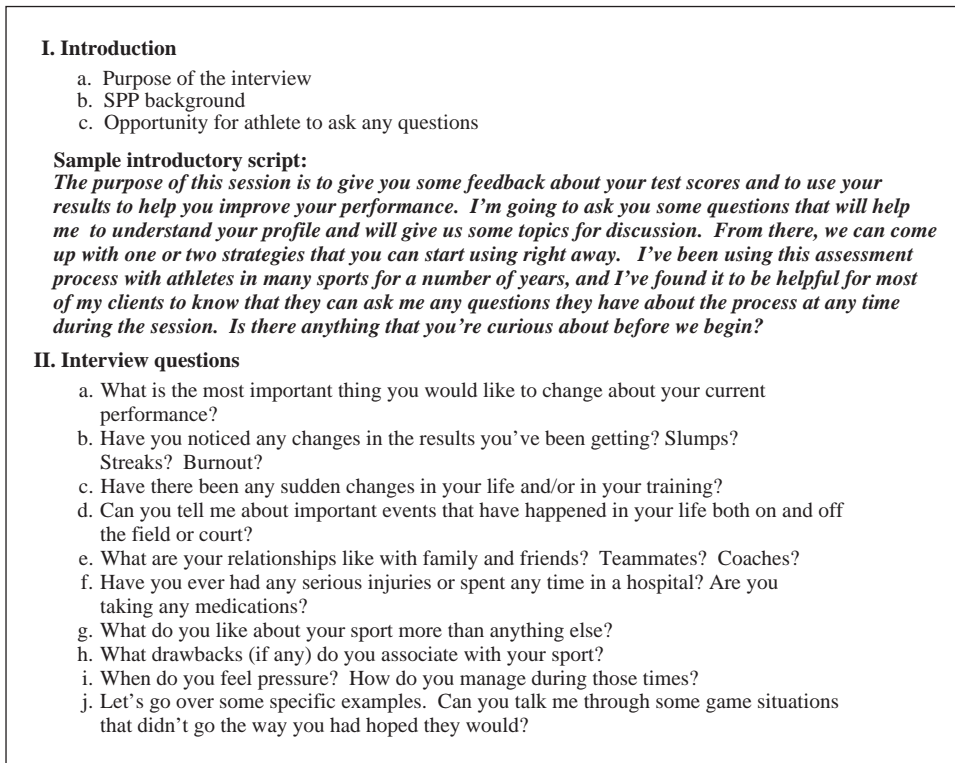

I. Introduction

    a. Purpose of the interview
    b. SPP background
    c. Opportunity for athlete to ask any questions

Sample introductory script:

The purpose of this session is to give you some feedback about your test scores and to use your results to help you improve your performance. I'm going to ask you some questions that will help me to understand your profile and will give us some topics for discussion. From there, we can come up with one or two strategies that you can start using right away. I've been using this assessment process with athletes in many sports for a number of years, and I've found it to be helpful for most of my clients to know that they can ask me any questions they have about the process at any time during the session. Is there anything that you're curious about before we begin?

II. Interview questions

    a. What is the most important thing you would like to change about your current performance?
    b. Have you noticed any changes in the results you've been getting? Slumps? Streaks? Burnout?
    c. Have there been any sudden changes in your life and/or in your training?
    d. Can you tell me about important events that have happened in your life both on and off the field or court?
    e. What are your relationships like with family and friends? Teammates? Coaches?
    f. Have you ever had any serious injuries or spent any time in a hospital? Are you taking any medications?
    g. What do you like about your sport more than anything else?
    h. What drawbacks (if any) do you associate with your sport?
    i. When do you feel pressure? How do you manage during those times?
    j. Let's go over some specific examples. Can you talk me through some game situations that didn't go the way you had hoped they would?

FIGURE 1   Sample interview outline.

of the issue(s), and the objective of the work. Many items on sport psychology inventories had their origins as interview questions. This should not be surprising. Like physicians who use intake forms and a wide range of tests to help learn about their patients, so too are SPPs interested in providing an efficient and consistent way in which to get to know their clients (Fig. 1).

3.2. Behavioral Observation

Behavioral observation is a straightforward process that involves watching live or recorded performance. It provides at least one distinct advantage over all of the other assessment methods in that SPPs can see firsthand how athletes behave in the most relevant settings (e.g., practice, competition). This information is invaluable because there are cues that may be identified in real-life settings that might not arise under less realistic conditions. Athletes might consistently make mistakes in situations but be unaware of the cause. The purpose of behavior observation is to watch athletes perform in those situations where insight into psychological performance factors can be ascertained. Both structured (i.e.,

controlled) and unstructured observation can be used to acquire this type of information. Trained SPPs can learn a great deal by attending to a variety of "targets" that include, but are not limited to, actual performance, verbal behavior, and body language. For example, slumped shoulders after a missed shot might reflect a slip in confidence. These kinds of observations can provide invaluable discussion points that can be critical in helping athletes to understand themselves and their performances (Table II).

TABLE II
Behavioral Observation Checklist

- *Verbalizations and utterances:* Listen for content, intensity, and timing.
- *Attention and nonverbal expressions:* Look for facial reactions, eye movement, and frequency, intensity, and general indicators of stress.
- *Posture and movement:* Look for bracing (stiffness), fluidity, limping, or other awkward positioning.
- *Social:* Observe interactions with teammates, parents, coaches, and the like.

3.3. Psychological Testing, Surveys, and Inventories

Although the definition of assessment offered here is rather broad, most people identify assessment with the printed questionnaire. Ranging from several questions to several hundred questions, these instruments usually ask athletes to respond to items by selecting from some form of multiple-choice response list. It should be noted that many SPPs are careful not to use the label "test" and prefer instead to use a variety of more "acceptable" names (e.g., survey, inventory, assessment). There are many ways in which instruments may be categorized, with some being more theoretically sophisticated than others. The following labels are commonly applied:

- Clinical
- Cognitive
- Attentional
- Personality and emotional intelligence
- Sport-specific
- Team/Social interactive

Although these different kinds of information are not uniquely attainable through psychological testing, the advantage of such testing is that it provides a more objective, more scientific means of acquiring data. Presumably, SPPs want to be able to know as much as possible so as to improve performance. Testing can provide a scientific psychological portrait of the athletes with whom they work. Clinical inventories can inform SPPs about qualities such as depression, anxiety, bipolarity, and mania. Cognitive inventories measure things such as verbal intelligence and speed of learning. Attentional inventories identify preferred styles of concentration. Personality tests involve measures of confidence, self-esteem, social attitudes, and risk taking. Sport-specific tools provide information about performance-relevant variables in sport settings. Team and social interactive inventories can help to identify factors particularly relevant in group settings. Examples would include measures of cohesion and communication style.

3.4. Third-Party Anecdotal Data Collection

Important information that is helpful to SPPs is often acquired through formal and informal communication with people in a special position to provide insight about clients. Third-party data collection may be as simple as an interview conducted by someone other than the SPP (and, as such, may be viewed as a type of interview), but it can also be much less formally acquired information by parents, teammates, coaches, or other individuals who come in contact with the athlete. One formal version of this often-overlooked method of assessment is similar to the 360-degree review process widely used in business consulting, where an individual receives feedback on performance strengths and weaknesses from his or her superiors, peers, and subordinates so as to help formulate an individual action plan for career development. Informal versions include casual conversations and other means of information exchange.

4. COMMON DISTINCTIONS

4.1. Formal Assessment Versus Informal Assessment

Behavior can be studied formally or informally. An informal assessment is one that does not incorporate psychometrics or data collected for statistical analysis. Examples of informal assessment include reading nonverbal cues when speaking with someone, watching sport competitions, observing practice sessions without recording data, and interviewing athletes, coaches, and parents. Formal assessment involves data collection through psychometric tools, surveys, frequency distributions, and the like. Discussion of assessment in sport psychology is dominated by formal topics because it is relatively easy to examine the rationale, experimental design, reliability, validity, and motivation for construction of formal instruments. Informal assessment is, on the other hand, undervalued and less stable due to the difficulty in standardizing techniques. But veterans in the field develop a balance of formal assessment (which provides a scientific base for their work) and informal assessment (which allows them to anticipate client needs and come up with practical solutions to identified problems as they happen in real time). It should be noted that although this distinction is commonly accepted, it is not always so clear. For instance, controlled observations (not usually viewed as formal assessment) can incorporate observation schedules with preestablished behavior categories and, thus, provide data eligible for statistical analysis.

There is both an art and a science to assessment. It is certainly worth noting that there are a fair number of "gurus" in the field who work primarily by intuition

and feel. Many respected SPPs rely most heavily on this "sixth sense," and it is likely that assessment by gut feelings will continue into the future.

4.2. Subjective Assessment Versus Objective Assessment

The distinction between subjective assessment and objective assessment is one commonly encountered in discussions of sport psychology. In general, subjective assessment encompasses various forms of interviews and observations on the part of SPPs. Objective assessment typically refers to psychological tools such as inventories and tests. Here, the distinction seems to hinge on the role of SPPs in the assessment process. In subjective assessment, SPPs are called on to interpret behavior directly. In objective assessment, SPPs use an appropriate instrument to provide additional information or clarification relevant to assisting athletes. What tends to confuse this distinction is that there are subjective components within objective assessment as well as objective components within subjective assessment. For example, formal scripted interviewing includes a measure of objectivity, and the interpretation of psychological inventories contains a large subjective component.

4.3. State Versus Trait

Because SPPs are ultimately interested in modifying behavior to improve performance, it is critical to understand whether a particular individual characteristic is likely to change over time. Traits are enduring characteristics that are likely present in a variety of situations. States are temporary conditions that are determined by situational factors. For example, even the most confident people can experience anxiety before an important speech or a big game. This could be called "state anxiety" and would be dependent on the situation in which an individual found himself or herself. However, an anxious person (i.e., one with "trait anxiety") would likely become anxious before every speech or game. Gauging where certain characteristics fall on the state–trait continuum is helpful to SPPs as they use assessment information and apply methods of intervention.

4.4. Sport-Specific Measures Versus General Measures

Some assessment tools contain items that are framed in a sport-specific context to elicit responses more relevant to the actual performance environment of the individual test-taker. Believing that sport-specific measures provide uniquely valuable information, many test developers have created sport-specific versions of existing inventories. This raises an interesting question: Is it more or less useful for items to focus on behavior within the situational context of primary concern? Not surprisingly, there is disagreement on the answer. Some believe that important information can be missed with sport-specific measures and that generally framed items will be more likely to pick up on those factors that are important in a variety of situations, including sport. This camp also believes that sport-specific items are more transparent and, thus, more easily faked or manipulated. Those who embrace sport-specific measures believe that the closer the context matches the particular performance environment of the athlete, the more validity the test will have and ultimately the more useful the results will be.

5. HISTORY AND CURRENT ASSESSMENT TOOLS (NORTH AMERICAN)

Assessment has played an integral role in the history of sport psychology. Most historical reviews identify the Triplett social facilitation experiments in 1898 as the first recorded studies in the field. In showing that cyclists raced faster with a pacer than they did by themselves, Norman Triplett used a simple measurement tool still believed by many to be the most critical component in athletic competition—results.

Nearly 30 years passed before a definitive text on the subject of athletes and their thought processes would appear. In 1926, Griffith published the first of two books that would become the "big bang" in the creation of sport psychology. *Psychology of Coaching* and *Psychology and Athletics* were written by Griffith as collections of his findings from the first ever sport psychology laboratory on the campus of the University of Illinois. In *Psychology of Coaching*, Griffith attempted to identify character types of coaches and men. His typologies included the "crabber," the "silent type," the "talker," the "loafer," the "leader," and the "follower."

The 1960s were a watershed for assessment and sport psychology. The first modern inventory created specifically for use with athletes was created in 1969 by Tutko, Ogilvie, and their colleagues. The Athletic Motivation Inventory (AMI) was born out of Tutko and Ogilvie's *Problem Athletes and How to Handle Them*. Until

this time, application of assessment in sport settings had meant adapting results of general personality and mental health diagnostics such as the Minnesota Multiphasic Personality Inventory (MMPI), 16 Personality Factor Questionnaire (16PF), and Thematic Apperception Tests. AMI attempted to measure traits categorized under the rubric of "Desire" factors and "Emotional" factors. Desire factors were drive, determination, aggression, leadership, and organization. Emotional factors were coachability, emotionality, self-confidence, mental toughness, responsibility, trust, and conscience development.

When international organizations began forming, communication increased between SPPs and assessment became a central focus. Could anxiety, personality, attention, moods, and motivation be tangibly measured, and could results be reliably reproduced, over time? If so, how would these tools contribute to the evolution of professional and amateur sport cultures in the real world? SPPs set to work on developing an arsenal of assessment tools for the sport environment.

One significant contributor, Rainer Martens, was a wrestling coach at the University of Montana who began to explore sport psychology after failing to help his talented team captain overcome prematch anxiety. In 1977, he published the Sport Competition Anxiety Test (SCAT) and, with help from colleagues in 1980, the Competitive State Anxiety Inventory (CSAI). Martens' work was preceded by Spielberger's theory that anxiety was present in both state and trait forms. Spielberger's State–Trait Anxiety Inventory (STAI) was published in 1970.

Meanwhile, Robert Nideffer's strong interest in aikido, as well as his expertise in MMPI research, led him to conclude that attention was the key psychological factor in sport performance. He developed The Attentional and Interpersonal Style (TAIS) Inventory while at the University of Rochester in 1976. His assessment tool directly measured the attentional preferences of athletes and helped practitioners to understand the conditions in which their clients would make concentration mistakes. In a survey taken 22 years later, only one other inventory, the Profile of Mood States (POMS), was being used more than TAIS by applied SPPs. POMS is a general inventory initially created in 1971 for use in psychiatry populations. Its popularity stems from its ability to assess both current moods and dominant mood traits.

The current state of assessment in sport psychology is one of many options. A directory of sport psychology assessments published by Ostrow in 1996 included 314 different inventories, questionnaires, and surveys. Sport psychology clients now include more than just elite athletes; they include a population of serious recreational athletes and aspiring juniors in many sports. Assessment tools such as Jesse Llobet's Winning Profile Athlete Inventory (WPAI) have been created to support developing athletes as well as professionals. Many assessments are now available on the Internet, where they can be used by athletes and practitioners from anywhere in the world in the privacy of their own homes. In professional ranks, player development and scouting departments count on psychological assessments and sport psychology consultants to help them in their coaching and drafting efforts. Examples of this application of psychological assessment include the use of AMI by the Major League Baseball Scouting Bureau and the use of Wonderlic by the National Football League (NFL) to assess intelligence levels of draft-eligible college players at the annual scouting combine. In addition, many professional teams rely on information from experts with self-designed assessments such as Robert Troutwine's Troutwine Athlete Profile (TAP).

Biddle's *European Perspectives on Exercise and Sport Psychology* provides a greater international perspective on history and current assessment tools.

6. LIST AND DESCRIPTION OF FREQUENTLY USED TESTS

This section contains descriptions of frequently used assessments employed in the field of sport psychology. The instruments are presented alphabetically by instrument name. Table III identifies the kinds of factors or variables measured by each assessment.

6.1. 16 Personality Factor Questionnaire (16PF)

Year created: 1950
Author: Cattell
Cattell believed that human personality consisted of 46 surface traits, from which 16 primary traits can be derived. Used primarily for sport psychological research, several studies showed meaningful differences on several scales for athletes and a significant correlation between extraversion and sport success. The 16 dimensions are Warmth (cool vs warm), Intelligence (concrete thinking vs abstract thinking), Emotional Stability (easily upset vs calm), Dominance (not assertive vs dominant), Impulsiveness (sober vs enthusiastic), Conformity (expedient vs conscientious), Boldness (shy vs venturesome),

TABLE III
Assessment Content Categories

Assessment [a]	Attention	Cognitive	Personality	Sport-specific	Team/Social interactive
16PF			X		X
ACSI			X	X	
AMI		X	X	X	
Competition Reflections		X		X	X
CSAI-2			X	X	
MBTI			X		
MMPI		X	X		
POMS			X		
PSIS		X	X		X
SCAT			X	X	
STAI			X		
TAIS (AME)	X		X	X	
TAP			X	X	
TOPS	X	X			
WPAI			X	X	X
WPT		X			

[a]16PF, 16 Personality Factor Questionnaire; ACSI, Athletic Coping Skills Inventory; AMI, Athletic Motivation Inventory; CSAI-2, Competitive State Anxiety Inventory; MBTI, Myers–Briggs Type Indicator; MMPI, Minnesota Multiphasic Personality Inventory; POMS, Profile of Mood States; PSIS, Psychological Skills Inventory for Sports; SCAT, Sport Competition Anxiety Test; STAI, State–Trait Anxiety Inventories; TAIS (AME), The Attentional and Interpersonal Style Inventory (Athletes Mental Edge); TAP, Troutwine Athletic Profile; TOPS, Test of Performance Strategies; WPAI, Winning Profile Athletic Inventory; WPT, Wonderlic Personnel Test.

Sensitivity (toughminded vs sensitive), Suspiciousness (trusting vs suspicious), Imagination (practical vs imaginative), Shrewdness (forthright vs shrewd), Insecurity (self-assured vs self-doubting), Radicalism (conservative vs experimenting), Self-Sufficiency (group-oriented vs self-sufficient), Self-Discipline (undisciplined vs self-disciplined), and Tension (relaxed vs tense).

6.2. Athletic Coping Skills Inventory (ACSI-28)

Year created: 1995
Authors: Smith, Schutz, Smoll, & Ptacek
This is a 28-item inventory designed with the intent to understand how athletes deal with competitive stress. ACSI-28 measures seven sport-specific subscales and then provides an overall "power" ranking that is a composite of those scale scores. The power ranking is called the Personal Coping Resources score, and the subscales are Coping with Adversity, Peaking under Pressure, Goal Setting/Mental Preparation, Concentration, Freedom from Worry, Confidence and Achievement Motivation, and Coachability.

6.3. Athletic Motivation Inventory (AMI)

Year created: 1969
Authors: Tutko, Lyon, & Ogilvie
This is a 190-item, forced-choice inventory that is historically the most widely referenced in the field of sport psychology. It was developed specifically to measure traits related to elite athletic success. Traits measured are grouped as Desire factors or Emotional factors. Desire factors are drive, determination, aggression, leadership, and organization. Emotional factors are coachability, emotionality, self-confidence, mental toughness, responsibility, trust, and conscience development.

6.4. Competition Reflections

Year created: 1986
Author: Orlick
This is composed of 12 open-ended items that provide opportunities for athletes to reflect on their "best" and "worst" selves by recalling details about great and poor performances. Included within these items are four

10-point rating scales that provide some reference for perceived arousal and anxiety. Direction of the reflections leads to the athlete brainstorming changes in training, preparing for competition, and interacting with coaches.

6.5. Competitive State Anxiety Inventory (CSAI-2)

Year created: 1990
Authors: Martens, Burton, Vealey, Bump, & Smith
Originally born out of Martens' work with SCAT, this is a modification of Spielberger's State Anxiety Inventory (SAI). The original version, created in 1980, isolated 10 items from SAI that measured competitive anxiety. A multidimensional version of the original, CSAI-2, includes three 9-item subscales measuring cognitive state anxiety, somatic state anxiety, and state self-confidence. In applied settings, a "Mental Readiness" form has derived from CSAI-2 that allows athletes to put themselves through a more convenient assessment of their anxiety levels while on the field or court.

6.6. Myers–Briggs Type Indicator (MBTI)

Year created: 1962
Authors: Myers & Briggs
Based on Jungian personality theory, this is the most widely used personality inventory in the world, with more than 2 million test takers annually. MBTI has a variety of applications in the worlds of personal development and management consulting. Its primary use in sport settings is for team cohesion and increasing awareness between teammates and coaches. Personality preference is measured along four dichotomies: Extraversion/Introversion, Sensing/Intuition, Thinking/Feeling, and Judging/Perceiving. Combinations of these scales produce a four-letter acronym that reflects the dominant score on each factor. These acronyms correspond to 16 different personality types that describe persons in detail according to behaviors associated with their personalities.

6.7. Minnesota Multiphasic Personality Inventory (MMPI-2)

Year created: 1942
Authors: Hathaway & McKinley
This is one of the cornerstones of psychiatric and clinical psychology assessment. MMPI is used to identify personality and psychosocial disorders as well as for general cognitive evaluation in adults and adolescents. A number of versions of MMPI exist, including the most recent version, MMPI-2, MMPI-A (an adolescent version), and short and long forms. Although MMPI is widely used in clinical settings, its use is limited in current applied sport psychology settings.

6.8. Profile of Mood States (POMS)

Year created: 1971
Authors: McNair, Lorr, & Droppleman
This is a 65-item instrument assessing tension, depression, anger, vigor, fatigue, and confusion. A 5-point scale (from "extremely" to "not at all") is used to rate responses. This inventory can be used to measure both state- and traitlike factors, depending on the "set" provided to the test taker by the proctor.

6.9. Psychological Skills Inventory for Sports (PSIS)

Year created: 1987
Authors: Mahoney, Gabriel, & Perkins
Consisting of 45 items with 5-point Likert response options, this instrument measures Anxiety Control, Concentration, Confidence, Mental Preparation, Motivation, and Team Emphasis. PSIS was created with input from 16 leading sport psychologists who answered the inventory questions as they thought the "ideal" athlete would. The original inventory asked true/false questions and contained 51 items. Test construction involved 713 athletes from 23 different sports.

6.10. Sport Competition Anxiety Test (SCAT)

Year created: 1977
Author: Martens
This is a 15-item inventory that was created to measure "state" anxiety levels in people anticipating competition. Martens was very clear in the original publication of the assessment that SCAT was not to be used to predict performance outcomes even if there was a relation to anxiety levels. Therefore, functional use of SCAT is optimal when gauging manifestations of anxiety in athletes during pregame or prematch periods.

6.11. State–Trait Anxiety Inventories (STAI)

Year created: 1970

Authors: Spielberger, Gorsuch, & Lushene

This was originally developed as a research tool for measuring anxiety in adults and has contributed mightily to sport anxiety research in a variety of populations. STAI is composed of two 20-item scales: one measuring State Anxiety and one measuring Trait Anxiety. The context for the A-State section is how the test taker is feeling at the current time, whereas the A-Trait section is based on general feelings of anxiety. Application of STAI has led to many significant findings in athlete work. For example, Yuri Hanin used STAI to create his Zones of Optimal Functioning (ZOF) model.

6.12. The Attentional and Interpersonal Style (TAIS) Inventory

Year created: 1976

Author: Nideffer

This is a 144-item inventory with an emphasis on concentration and the critical interpersonal factors affecting performance under pressure. The latest version of TAIS is composed of 20 scales that range principally over attentional and interpersonal factors. The two newest scales are Focus over Time and Performance under Pressure. Athletes Mental Edge (AME) used a 50-item version of TAIS and was introduced in 1999. AME was the first widely used, Internet-based sport psychological assessment tool.

6.13. Troutwine Athletic Profile (TAP)

Year created: 1988

Author: Troutwine

This is a 75-item questionnaire that receives considerable attention prior to the NFL draft. Troutwine does not publish his assessment for distribution, so there are few SPPs using TAP. However, this may be the most well-known, "self-written for self-use" assessment in the field. The questionnaire generates a Player Report for development and a Coach Report for recruiting and selection; this is why many NFL teams privately contract with Troutwine and value his data. The Player Report measures Competitive Nature, Coping Style, Work Style, and Mental Characteristics. The Coach Report topics are Mental Makeup, Social Style, Personal Characteristics, Coaching Suggestions, and Performance Tendencies.

6.14. Test of Performance Strategies (TOPS)

Year created: 1999

Authors: Thomas, Murphy, & Hardy

This is a 64-item inventory intended for use when measuring psychological skills and strategies used by athletes both in competition and while at practice. TOPS results differentiate between practice and competitive skills and strategies on the following scales: Goal Setting, Use of Imagery, Ability to Relax, Ability to Activate/Energize, Use of Positive Self-Talk, Emotional Control, Attentional Control, Automaticity (i.e., performing without thinking), and Negative Thinking. Although TOPS is a fairly new inventory, its applied use was already seen with U.S. and Australian Olympic teams and coaches in preparation for the Sydney Games in 2000.

6.15. Winning Profile Athlete Inventory (WPAI)

Year created: 1999

Author: Llobet

This is composed of 55 items with 5-point Likert response options that measure a wide variety of psychological skills in athletes. WPAI responses load onto 10 factors: Competitiveness, Dependability, Commitment, Positive Attitude, Self-Confidence, Planning, Aggressiveness, Team Orientation, Willingness to Sacrifice Injury, and Trust. Functional use is intended for athletes from high school to professional levels for recruiting and development purposes.

6.16. Wonderlic Personnel Test (WPT)

Year created: 1937

Author: Wonderlic

This is a 50-item test of general intelligence that the user has 12 minutes to complete. Originally designed for use with job applicants in both blue- and white-collar settings, WPT has become a mainstay in the assessment of collegiate football players at the NFL's scouting combine. The test has 16 different versions with problems that include algebra, word analogies, mathematical word problems, and various other tests of verbal, computational, and analytic abilities. It measures an individual's ability to learn, adapt, solve problems, and understand instructions.

7. FORMAL ASSESSMENT GUIDELINES AND ETHICAL CONSIDERATIONS

When done well, psychological testing provides professionals with an efficient, objective, and repeatable approach to assessing behavior. Many factors affect the quality of the information gathered through the various assessment means. Sensitivity to issues around test validity, confidentiality, response sets, and response styles is critical for effective use of sport psychology assessment.

7.1. Test Validity

It is incumbent on SPPs to evaluate the validity of the assessment methods they use. A test's validity determines how accurately it measures what it purports to measure. There are three commonly referenced types of validity: face, construct, and predictive. Face validity is a surface measure that acts as a first step toward the integrity of an assessment. Face validity is attained when the parameters of a test seem like they would generate the results that SPPs hope to generate. Construct validity links theory with practice. It can be summed up as the ability that an assessment has to demonstrate, with results, the topic it is intended to measure. Predictive validity is simply the likelihood that the results of an assessment can predict the occurrence of a specified behavior in the future.

7.2. Confidentiality

According to the code of ethics set forth by the American Psychological Association, the confidentiality of the psychologist–client relationship is protected. Because many nonlicensed individuals function as sport psychology consultants, the status of such a relationship is less clear and not likely afforded the same kind of protection. When it comes to the confidentiality of assessment results, SPPs should be careful to protect confidentiality explicitly and to explain clearly to clients who will have access to their assessment information. Because it can sometimes be difficult to identify exactly who the client is (e.g., the athlete, the coach, the organization), it is advisable to have a written policy that describes how assessment information will be kept and shared. Many SPPs use informed consent forms to establish clearly the confidentiality of assessment data (Fig. 2).

7.3. Response Sets and Styles

When asking athletes to answer any set of questions about themselves, it is important to take into account variables that will influence the accuracy of their responses. Response sets are situation-specific attitudes that individuals adopt when completing a psychological assessment. They are often employed when athletes are skeptical of how their assessment data may be used. Faking can occur in situations where athletes feel pressure to present themselves in an ideal light.

In contrast to response sets, response styles reflect cross-situational, core personality characteristics. Common response styles include overly conservative and extreme appraisals. People who adopt a conservative response style see the world in shades of gray and tend to respond with middle-range answers such as "maybe" and "sometimes" on inventories with Likert

---

Assessment Consent Form

I hereby grant permission to _____ for sport psychology assessment. I understand that any information gained from testing will be used for the purpose of performance enhancement and will be kept confidential. I further understand that none of my personal data will be released to anyone without my consent. I agree that although my individual data may be used as part of a database for research, my identity and other personal information will not be disclosed.

I understand that my access to sport psychology services is not contingent on my granting of this permission. In addition, I understand that if at any time during the assessment phase I feel uncomfortable, I can end the assessment process.

_____ _____
Signature   Date

_____ _____
Printed name   Sport psychology professional

FIGURE 2　Sample informed consent form.

scale choices. Extreme response styles are often seen in individuals who have either an overly optimistic outlook or an overly critical outlook. Aside from the impact that response sets and styles have on individual scores, SPPs must be careful when comparing different athletes' assessment results. Although some degree of influence from both response sets and response styles is always likely to be present, their effects can be minimized by encouraging open and honest answering and by providing points of comparison for various response choices. Many SPPs prefer not to control for response sets because the sets and styles themselves provide a great deal of information about the athletes.

7.4. Consensual Validation

To ensure that test information is accurate, it is important to have the opportunity to speak directly with the individuals who have completed the assessments to determine the impact of any existing response sets or response styles.

Consensual validation is the process of verifying the interpretations made from test information, which typically occurs through interview or direct observation. Effective consensual validation requires the ability to define operationally psychological constructs (e.g., confidence, control). By translating these concepts into specific behaviors, SPPs can determine their presence or absence more accurately. Imagine that a client completes a personality inventory and scores very high on control and self-esteem. A high score on these dimensions would lead the interpreting SPP to assume certain behaviors in various circumstances. More often than not, the SPP would expect the client to speak up, take charge, be reluctant to listen to contrary opinions, and oppose authority. The SPP would be looking to verify that these behaviors occur with the kind of frequency commensurate with the scores on the performance-relevant construct(s). Table IV offers a number of suggestions to increase effective consensual validation.

8. FUTURE DIRECTIONS

Just as athletes engage in the "pursuit of perfection," SPPs seek the most effective means of helping the athletes with whom they work. Technology is already having a pronounced effect on how assessment is carried out. Psychological testing can now take place over the Internet with the tester and the testee thousands of miles apart. More and more individuals will gain access to assessment information. Feedback

TABLE IV
Consensual Validation Techniques

- Select tests that measure constructs that are easily translatable into observable behaviors.
- Select tests that measure a number of performance-relevant variables or choose several tests that measure a variety of constructs. The reason for this is that individual constructs are rarely able to explain performance across individuals and situations. A number of distinct cognitive and interpersonal characteristics must be examined to effectively account for the influences that individual and situational differences have on performance.
- Translate each construct into expected behaviors (for the performance arena in which they operate).
- Create conditions under which the individual's behavior can be observed under varying degrees of pressure. Remember that most people are able to control even dominant behaviors when they are not experiencing significant stress. The key is to see how much difference increasing the pressure has on the presence or absence of particular behaviors.

Source. Adapted from Nideffer and Sagal (2001).

reports delivered instantaneously, and sometimes without professional interpretation, will most certainly force the field of sport psychology to confront what is and is not acceptable practice.

As athletes are exposed to the benefits of working with SPPs, it is inevitable that more athletes themselves will develop a professional interest in the field. Consequences of this are likely to include the following:

- An increased respect for the field of sport psychology itself
- A greater emphasis on practical applications in specific sports
- Increased pressure on SPPs to have significant sport-specific athletic experience of their own

Although the focus of this article has been on assessment in applied sport psychology, it is important to note that it is nearly always the case that application follows research. Advancements in medicine and new insight into the mind–body connection will likely increase the use of neurological, chemical, and physiological testing to help predict and influence performance. The ongoing search for strong ties between emotions and body chemistry may bring new kinds of SPPs to the field. For example, heart rate monitors and other biofeedback mechanisms are already being used to help athletes achieve optimal levels of arousal and to help SPPs better understand the competitive experience (Fig. 3).

1913: Fifth International
Olympic Conference held at
Lausanne, Switzerland
(topic: Sport Psychology
and Physiology)

1926 and 1928: Griffith publishes
landmark titles:
Psychology of Coaching and
Psychology and Athletes

| 1900 | 1905 | 1910 | 1915 | 1920 | 1925 | 1930 | 1935 | 1940 | 1945 | 1950 |

1898: Triplett conducts
social facilitation experiments;
first record of sport psychology

1938: Chicago
Cubs hire Griffith as
their sport
psychologist

FIGURE 3 Assessment timeline.

1965: International Society of Sport Psychology (ISSP) is formed

1976: Nideffer develops **The Attentional and Interpersonal Style (TAIS) Inventory**

1986: American Psychological Association approves Division 47: "Exercise and Sport Psychology"

1986: Vealey develops **Trait and State Sport Confidence Inventories**

1958: Brazil wins World Cup with Joao Carvalhaes on staff as team psychologist

1977: Martens develops **Sport Competition Anxiety Test (SCAT)**

1986: Orlick publishes *Psyching for Sport: Mental Training for Athletes,* which includes Competition Reflections instrument

1999: The Internet becomes a tool for sport psychology assessment; Nideffer and Sagal create the **Athlete's Mental Edge (AME)**

1966: North American Society for the Psychology of Sport and Physical Activity (NASPSPA) is formed

1955   1960   1965   1970   1975   1980   1985   1990   1995   2000

1966: Ogilvie and Tutko's *Problem Athletes and How to Handle Them* is published

1970: Spielberger, Gorsuch, and Lushene publish *Manual for the State–Trait Anxiety Inventory (STAI)*

1985: Association for the Advancement of Applied Sport Psychology (AAASP) is formed

1996: Ostrow's *Directory of Psychological Tests in the Sport and Exercise Sciences* lists 314 different inventories, questionnaires, and surveys used in sport

1969: Tutko, Lyon, and Ogilvie develop **Athletic Motivation Inventory (AMI)**

1971: McNair, Lorr, and Droppleman release **Profile of Mood States (POMS)**

1990: Martens, Burton, Vealey, Bump, and Smith create **Competitive State Anxiety Inventory-2 (CSAI-2)** as an update to original inventory
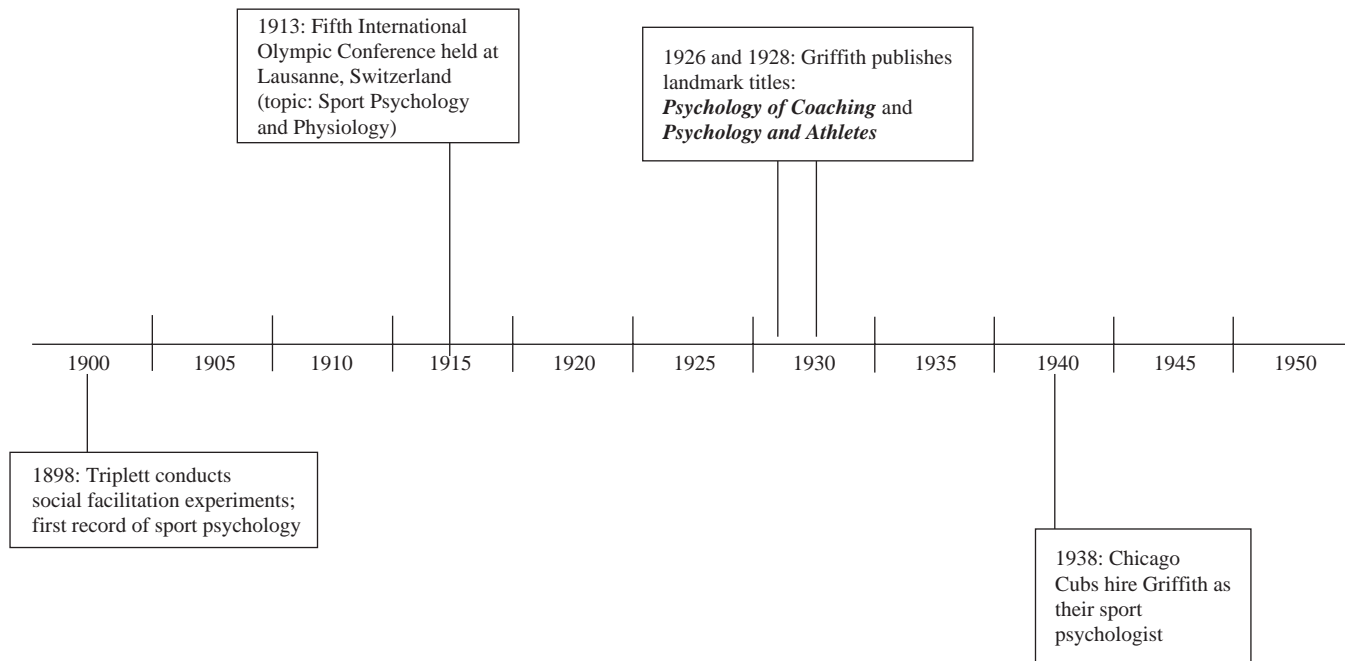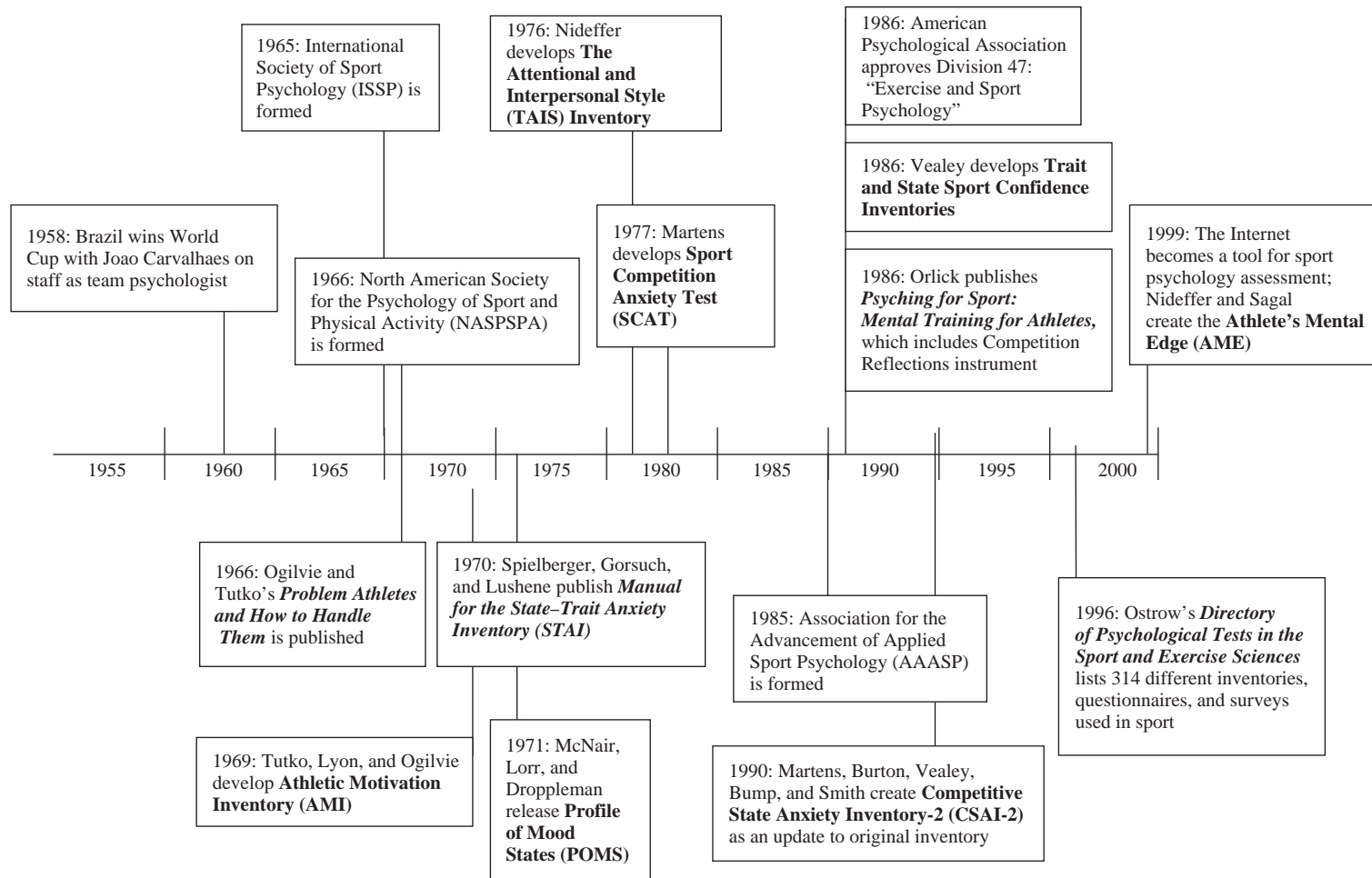
FIGURE 3 (*Continued*).

See Also the Following Articles

Anxiety and Optimal Athletic Performance ■ Attention and Concentration Training in Sport ■ Goal Setting and Achievement Motivation in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Performance Slumps in Sport: Prevention and Coping ■ Psychological Skills Training in Sport ■ Self-Confidence in Athletes

Further Reading

Anastasi, A. (1988). *Psychological testing*. New York: Macmillan.

Biddle, S. J. (Ed.). (1995). *European perspectives on exercise and sport psychology*. Champaign, IL: Human Kinetics.

Duda, J. L. (Ed.). (1998). *Advances in sport and exercise psychology measurement*. Morgantown, WV: Fitness Information Technology.

Fisher, A. C. (1984). New directions in sport personality research. In. J. M. Silva, & R. Weinberg (Eds.), *Psychological foundations of sport* (pp. 70–80). Champaign, IL: Human Kinetics.

Heil, J., & Henschen, K. (1996). Assessment in sport and exercise psychology. In J. Van Raalte, & B. Brewer (Eds.), *Exploring sport and exercise psychology* (pp. 229–255). Washington, DC: American Psychological Association.

Nideffer, R. M., & Sagal, M. S. (2001). *Assessment in sport psychology*. Morgantown, WV: Fitness Information Technology.

Orlick, T. (1990). *In pursuit of excellence* (2nd ed.). Champaign, IL: Human Kinetics.

Ostrow, A. C. (1996). *Directory of psychological tests in the sport and exercise sciences*. Morgantown, WV: Fitness Information Technology.

Scott, S. (2002). *Fierce conversations*. New York: Penguin Putnam.

Silva, J. (1984). Personality and sport performance: Controversy and challenge. In J. Silva, & R. Weinberg (Eds.), *Psychological foundations of sport* (pp. 59–69), Champaign, IL: Human Kinetics.

Attachment

Ronit Roth-Hanania and Maayan Davidov

University of Haifa, Haifa, Israel

GLOSSARY

anxious–avoidant attachment An insecure form of attachment; during the Strange Situation, infants classified as avoidant appear to be relatively indifferent to separation from the caregiver even though physiological markers indicate that they are in fact distressed, and on reunion, they actively ignore and avoid the caregiver, diverting their attention to exploration.

anxious–resistant (ambivalent) attachment An insecure form of attachment; during the Strange Situation, infants with this classification exhibit extreme distress on separation, and when the caregiver returns, they refuse to settle down, clinging while at the same time pushing the caregiver away in anger.

attachment An emotional connection to another individual that develops between an individual who is seen as able to provide protection, comfort, and support in times of need; early attachment develops between an infant and a primary caregiver through patterns of caregiving interactions that evolve over the first year of life.

disorganized attachment A Strange Situation classification given in addition to one of the other three classifications; infants classified as disorganized/disoriented show sequences of behavior that appear to lack a clear intent, purpose, or organization (e.g., freezing).

internal working models Internal mental representations that are established during the formation of an attachment relationship; these mental representations refer to both self and the attachment figure, and later in life they guide a person's processing of relevant information and behavior in the context of interactions and relationships.

secure attachment A positive and adaptive form of attachment between an infant and his or her caregiver; the infant uses the caregiver as a base for independent exploration of the environment and as a source of comfort and support during times of stress and distress.

Strange Situation procedure A systematic laboratory observational method designed to assess the quality of attachment organization of a 12- to 18-month-old infant to a caregiver.

Attachment theory, put forth by John Bowlby, postulates that infants develop strong emotional ties to consistent caregivers during the first year of life and that these ties are crucial for the individual's normal social–emotional development and well-being. Bowlby, a British psychiatrist, combined ideas from psychoanalysis, ethology, and evolution to develop a theory regarding the attachment behavioral system—its nature, function, and development. Mary Ainsworth, Bowlby's collaborator, provided empirical support for many of Bowlby's ideas in her in-depth observations of

mother–infant interactions at home and in a laboratory environment. Ainsworth also developed a laboratory observation procedure, the Strange Situation, which enabled her to identify individual differences in attachment organization among 1-year-old infants. The Strange Situation procedure, currently the most widely used and validated measure of attachment security during infancy, is composed of a series of short separations and reunions between an infant and his or her caregiver, usually the mother. It enables psychologists to classify the quality of an infant's attachment to a caregiver into one of three major categories: secure attachment, anxious–avoidant attachment, or anxious–resistant attachment. A fourth category, disorganized/disoriented attachment, was added more recently. Ainsworth and colleagues also demonstrated that the caregiver's sensitive responsiveness to the infant's signals and cues is important for the formation of secure attachment. This finding has been replicated in many subsequent studies, although not always so strongly. Research on attachment yields some interesting findings regarding the inter-generational transmission of attachment, the stability of attachment throughout the life span, and the role of early attachment in later child outcomes, particularly in the areas of emotion regulation, cognition, social adjustment, and mental health. Because early attachment relationships are viewed as crucial to later development, there is a growing number of prevention and intervention programs designed to promote attachment security among populations at risk for insecure attachment relationships. The effectiveness of these programs has varied. Ever since it was first introduced, attachment theory has played a central role in shaping our understanding of child development and parent–child relations, as reflected in developmental theory and research, clinical and applied settings, and policymaking.

1. INTRODUCTION

Attachment theory centers on the notion that emotionally responsive care, including love and nurturance from a primary caregiver, is essential for healthy and normal development. John Bowlby, the founder of attachment theory, used the term "attachment" to describe the emotional connection that develops between an infant and a primary caregiver through patterns of interaction that evolve over time. According to attachment theory, during the first year of life, infants develop special ties to their primary caregivers (i.e., attachment figures). Every infant invariably becomes

attached unless no single caregiver is continuously present to care for the child (e.g., as in some institutions and orphanages). The particular nature or quality of the attachment relationship will vary, depending on the history of interaction patterns within the dyad. In particular, the consistency and appropriateness of the caregiver's response to the infant during times when the infant feels stressed or threatened, defines the pattern of interactions that develop during the first year. When the caregiver is consistently sensitive and responsive, a secure attachment typically ensues, whereas insensitive or inappropriate responding usually leads to an insecure attachment relationship.

After the attachment is formed, the infant shows a preference for the attachment figure. This is particularly evident during times of stress or emotional upset. When scared or distressed, the infant actively seeks comfort and reassurance from the caregiver and is not easily comforted by other individuals. The presence of the attachment figure is reassuring and enables more competent exploration of the environment. Mary Ainsworth, Bowlby's collaborator, noted that in a secure attachment relationship, the attachment figure provides both a "safe haven" during times of need and a "secure base" from which to explore.

Bowlby viewed the attachment system as a biobehavioral system that organizes behavior. The set goal of this system is to maintain proper access and proximity to the caregiver during times of need to ensure the protection and survival of the young. The attachment system operates in interaction and in collaboration with other behavioral systems (e.g., fear, exploration, sociability).

Attachment relationships are believed to play an important role throughout the life cycle. Although the specific behavioral markers of attachment change, their function remains the same. During times of need, infants, older children, and adults alike all seek comfort and support from significant others who are seen as more capable of coping with these situations. As individuals mature, new attachments are typically formed, but the early attachment relationships remain important because they are thought to exert an effect on subsequent behavior and close relationships.

2. HISTORY

2.1. John Bowlby (1907–1990)

Bowlby, a British child psychiatrist and psychoanalyst by training, was one of the first to systematically

observe the negative effect of impoverished family life and the lack of consistent care on the development of young children. He suggested, based on his work with hospitalized and homeless children during the 1940s, that a child needs to experience a close, warm, and consistent relationship with his or her mother or another permanent caregiver so as to develop normally.

Influenced by ethology and evolutionary theory, Bowlby rejected the then widely accepted notion that an infant's tie to his or her mother is rooted in the mother's role as a provider of food or oral gratification. Lorenz's observations of the "imprinting" phenomenon in geese during the 1930s demonstrated that an organism can become "attached" to its caregiver through a process unrelated to feeding. Harlow's laboratory studies of rhesus monkeys during the 1950s and 1960s also demonstrated that needs other than hunger are important in the formation of the infant–mother tie. For example, monkeys "raised" by two kinds of surrogate mothers—a wire mesh "mother" to which a bottle with milk was attached and a terrycloth surrogate "mother" with no milk bottle—preferred the soft cloth mother, spending most of their time clinging to her for comfort and rushing to her when frightened. They spent little time on the wire mother, approaching her only when wanting to be fed. Influenced by these and other studies, Bowlby posited that the need for proximity to a caregiver is a biologically based primary need that evolved to ensure the protection and survival of the young.

Bowlby also argued that it is children's actual experiences with caregivers (e.g., lengthy separations, lack of consistent care) that determine their well-being and adjustment. This view was at odds with the notions accepted by the psychoanalytic community, which placed greater emphasis on children's fantasy lives than on their actual experiences with caregivers.

These disagreements led Bowlby to establish his own research unit at the Tavistock Clinic in London, where he proceeded to develop his groundbreaking ideas. His work is summarized in the trilogy titled *Attachment and Loss,* published in 1969, 1973, and 1980, as well as in many other books and articles. The focus in Bowlby's writings was primarily on the evolutionary origin and function of the attachment system, how it organizes behavior, its developmental course, and the effects of early separations and loss of attachment figures on a person's well-being and adjustment. The work of Ainsworth provided empirical grounding for Bowlby's theoretical thinking, extended Bowlby's ideas on attachment, and brought to the forefront a focus on different patterns of attachment organization.

2.2. Mary Ainsworth (1913–1999)

Ainsworth met Bowlby during the early 1950s at the Tavistock Clinic. She subsequently carried out an observational study in Uganda designed to examine Bowlby's ideas empirically. After returning to the United States, she proceeded to conduct a second attachment study in Baltimore, Maryland, involving intensive home observations of 23 mother–infant dyads. As part of the Baltimore study, Ainsworth also developed a systematic laboratory observational method that enabled her to observe the reactions of infants to short sequences of separations from their mothers. This procedure, known as the Strange Situation, allowed her to identify individual differences in infants' attachment to their mothers, that is, their different strategies of using the mother when dealing with the stress brought about by brief separations. The Strange Situation procedure initially yielded three types or classifications of attachment: a secure pattern and two insecure classifications (anxious–avoidant and anxious–resistant attachment). The procedures and findings of the Baltimore study are summarized in the book *Patterns of Attachment,* published in 1978. Ainsworth's work generated a surge of empirical studies that employed the Strange Situation and focused on individual differences in attachment security.

3. MEASUREMENT

Over the years, researchers have developed a variety of measures to assess attachment patterns at various ages. The most widely used and best validated assessment procedures employed with infants, children, adolescents, and adults are described in what follows.

3.1. Assessing Infant Attachment

3.1.1. The Strange Situation

3.1.1.1. Procedure For more than 30 years, quality of parent–infant attachment has been assessed using the Strange Situation procedure developed by Ainsworth in the Baltimore study. This procedure, which is appropriate to use with infants ages 12 to 18 months, involves a series of moderately stressful separations (3 minutes each) and reunions (3 minutes each) between the infant and a caregiver while in an unfamiliar playroom. Two systems are activated during the Strange Situation procedure: exploration (of a new environment, toys, and people) and attachment (through the repeated separations

and reunions). The Strange Situation procedure enables researchers to examine how the infant uses the caregiver to cope with stress. The infant's strategy is thought to reflect his or her accumulated experience as it pertains to the caregiver's availability during times of need.

3.1.1.2. Attachment Patterns   Ainsworth identified three major attachment classifications. Secure infants use their caregivers as a "secure base" for exploration. These infants rely on occasional visual, verbal, or physical contact with their mothers as a basis for their own initiated exploration of the environment. When their mothers leave the room, secure infants may or may not cry. On reunion, these infants either greet their mothers positively or, if upset, go to them for comfort and then shortly after return to activities associated with exploration.

Infants who show patterns of insecure attachment of the avoidant type exhibit little or no reference to their mothers while exploring the room and show no overt signs of distress on their mothers' departure (although some studies have shown that these infants do show physiological markers of stress during separation). Most important, on reunion, these infants actively ignore and avoid their caregivers (e.g., by looking away).

Insecure–resistant (or ambivalent) infants are preoccupied with their mothers' presence, often unable to leave their sides even in light of curious attempts to explore their new environment. When their mothers leave the room, these infants become extremely distressed, and on reunion, the infants refuse to settle and to resume exploration, clinging to their mothers and at the same time expressing anger and dissatisfaction.

The large body of work using the Strange Situation shows that across cultures, the majority of all infants (approximately 65%) are securely attached to their mothers, whereas the remainder are classified as insecure. The distribution of insecure attachment classifications (i.e., frequencies of the avoidant and resistant patterns) appears to be affected by cultural context. In Western Europe and the United States, the avoidant type is more prevalent than the resistant pattern, whereas the opposite has been observed in some other cultures (e.g., in Israel).

A fourth category of attachment classification was added in 1986 by Main and Solomon because studies showed that there are infants whose behavior does not correspond with any of the three existing classifications. During the Strange Situation, these infants showed sequences of unusual behavior with no clear purpose or orientation (e.g., freezing, falling prone).

They seemed to lack an organized strategy for coping with the stress of the situation and looked "confused"; thus, they were classified as having disorganized/disoriented attachment to their caregivers. Disorganized classification is given in addition to one of the other classifications. In normative middle-class populations in North America, approximately 14% of infants are judged to be disorganized. The rate is somewhat higher (24%) in low-socioeconomic status samples.

3.1.1.3. Validity   Evidence supporting the validity of the Strange Situation (i.e., its capacity to indeed capture the organization of the child's attachment system) comes from the correspondence between Strange Situation classification and Ainsworth's home observations of the same 23 infant–mother pairs. Ainsworth and colleagues found a strong association between infants' Strange Situation classification and their secure base behavior in the home. Secure infants were much more likely than insecure infants to exhibit a good balance of exploratory and proximity-seeking behavior, to transition smoothly between the two, and to enjoy and be comforted by physical contact. In contrast, insecurely attached infants showed disturbed balance between attachment and exploration behavior, had difficulty in transitioning from one to the other, and often expressed negative affect during physical contact with their mothers. Further supporting evidence comes from several studies that found links between the Strange Situation classification and infants' secure base behavior in the home as assessed by the Attachment Q-sort. Finally, additional less direct evidence for the validity of the Strange Situation comes from its links to prior maternal behavior and subsequent child outcomes.

3.1.2. The Attachment Q-Sort

An alternative to the Strange Situation is the Attachment Q-sort developed by Waters and Deane in 1985. The AQS requires observers to describe a target child's home behavior by sorting a set of items according to how well they each describe the particular child (from most characteristic to least characteristic). Version 3.0 of the AQS consists of 90 items depicting a wide array of child home behaviors. A continuous measure of security is derived by comparing the sort describing the child of interest with a criterion sort developed to describe the ideal secure child. An advantage of the AQS is that it is suitable for use with a relatively wide age range (from 1

to 4 years). Both mothers and independent observers may use the AQS, although mothers have to become familiar with its items and observe their children for a few days in the home environment prior to their finally sorting all of the items from most characteristic to least characteristic of their children.

Findings regarding the association between the Strange Situation classifications and AQS security score have been mixed; some studies have found correspondence, whereas others have not. Overall, it appears that correspondence is better when independent observers, rather than mothers, are used. Summarizing more than 130 studies, a recent meta-analysis showed moderate convergence between observer-reported AQS security and Strange Situation security. Observer reports of AQS security were also found to be quite strongly linked to maternal sensitivity. This suggests that the AQS, when rated by observers, is a reasonable alternative to the Strange Situation. On the other hand, maternal-reported AQS security had weaker associations with both Strange Situation security and maternal sensitivity assessments, suggesting that mother-reported AQS scores are less adequate at assessing attachment security.

3.2. Measuring Attachment During Childhood

Beyond infancy, measurement of the attachment relationship relies heavily on observations of children's behavior on reunion following relatively long separations from parents (30–60 minutes). Children's emotional expression during separation and reunion, as well as their patterns of verbal and nonverbal communication with their caregivers, is examined and used to determine their attachment classification. The three main attachment patterns identified in children parallel those identified in the Strange Situation. A child can be described as secure (also referred to as balanced), as avoidant (also termed defended), or as dependent (parallel to the resistant classification during infancy and also referred to as coercive or involving). Children classified as disorganized during infancy typically exhibit controlling role-reversing behavior on reunion during childhood, taking either a caregiving or a punitive role toward their parents.

Narrative measures attempting to assess children's symbolic representations of attachment have also been developed. These measures use projective techniques. They ask children to respond to pictures or story-stems depicting relevant attachment scenes. The narratives constructed by the children are then coded and scored for attachment style. Finally, the AQS can also be used with young children to obtain a measure of attachment security.

3.3. Assessing Adult Attachment Representations

During adolescence and adulthood, attachment is most often assessed through structured interviews or questionnaires. The central instrument is the Adult Attachment Interview (AAI) developed by Main and colleagues. Throughout the interview, the individual is asked to describe his or her relationships with caregivers during childhood and to illustrate these descriptions with specific examples. The individual is also asked to describe situations where he or she needed support from those attachment figures (e.g., when the individual was hurt or upset) as well as relevant experiences of loss or trauma. The interviews must be transcribed verbatim and scored by a trained coder.

The AAI is designed to assess the individual's current state of mind with respect to attachment rather than what his or her early attachment classification might have been. Thus, scoring of the AAI is based on the coherence of the narrative produced by the individual during the interview rather than on an assessment of his or her early life experiences.

The AAI yields adult attachment classifications that mirror the patterns identified in the Strange Situation. Securely attached adults are referred to as "autonomous" or "free to evaluate" because they appreciate close relationships and value their effect. They are reflective and open, and they speak about early experiences in a coherent and objective manner, even when such experiences were difficult or painful. Insecure adults could be classified as either "dismissing" (parallel to avoidant during infancy) or "preoccupied" (parallel to resistant attachment during infancy). Dismissing individuals minimize the effect of meaningful relationships in their personal history. They may idealize or derogate their attachment figures. When asked to support their descriptions of their relationships with specific examples, they are often unable to remember relevant incidents. Preoccupied (or "enmeshed") individuals are unable to detach themselves from irrelevant details associated with their past history and are so wrapped up in their "old" experiences and feelings that they fail to provide an overview. They appear to still be immersed in anger over old issues and unable to "let go."

Finally, the coding of the AAI also yields a score of the individual's unresolved loss or trauma (when such painful experiences had occurred). This score is based on indications of disorganization and disorientation in thinking and affect during the discussion of the loss/trauma (e.g., speaking of a dead person as if he or she were alive, going into an extreme amount of detail when describing the events surrounding the loss).

AAI classifications have been found to be stable over time and unrelated to intelligence, general memory skills, or the individual's narrative style about topics irrelevant to attachment. Some empirical evidence regarding the validity of the AAI comes from numerous studies that found links between mothers' AAI classification and the quality of their caregiving behavior to their own children at home as well as their own children's attachment classification. More direct evidence that the AAI indeed reflects how adults interact with attachment figures when coping with stress is still relatively scarce. In one relevant study, engaged couples who completed the AAI were observed during a standard marital interaction task. AAI security was associated with more adaptive use of the partner as a secure base to obtain support as well as with greater propensity to serve as a secure base for the partner and provide such support. More research linking the AAI to adult secure base behavior is still needed.

Other interview procedures have also been developed to assess adult attachment. For example, the Current Relationship Interview (CRI) was developed by Crowell and Owens to assess attachment within current close relationships. Finally, self-report questionnaires are used as well for the assessment of attachment styles, particularly in the context of adult romantic relationships. Of these, the most widely used are Hazan and Shaver's Attachment Style Questionnaire and Bartholomew and Horowitz's Relationship Questionnaire. Overall, studies examining the associations between questionnaire measures and the AAI have produced mixed findings; several studies find little or no relationship, whereas a few studies report moderate associations. More research is still needed concerning the relationship among various measures of adult attachment.

4. THE ANTECEDENTS OF ATTACHMENT SECURITY

Because attachment describes an aspect of the relationship that develops between an infant and a caregiver,

both members of the dyad influence the nature and quality of the attachment relationship that emerges. These two classes of antecedents are reviewed in what follows. Genetic influences on attachment are also considered.

4.1. Caregiver Influences on Attachment Security

4.1.1. Parental Caregiving Behavior

In Ainsworth's studies, infant attachment classification was found to be strongly linked to prior maternal behavior in the home. The extensive home observations revealed that the mothers of secure, avoidant, and resistant infants differed qualitatively in their styles of caregiving. Specifically, mothers of infants later found to be secure during the Strange Situation at 12 months were highly sensitive, responsive, and accepting. They responded consistently and appropriately to their infants' signals (e.g., crying) and were affectionate when holding and taking care of their infants. In contrast, the mothers of avoidant infants were observed to be consistently insensitive and rejecting. They were interfering and showed an aversion to close physical contact with their infants as well as a relative lack of emotional expression. Finally, infants classified as resistant had mothers who were inconsistently responsive. They were generally insensitive to their infants' signals but were not highly rejecting. They were inept at holding their infants and responding to signals, but they showed no aversion to close physical contact with their infants. Therefore, the strategies employed by the insecure infants during the Strange Situation were seen as adaptations to the nonoptimal caregiving they have received and come to expect.

Overall, Ainsworth and colleagues found a very strong positive association between observer ratings of maternal sensitive responding and Strange Situation security. A considerable body of subsequent work employing the Strange Situation has generally replicated this link between early maternal sensitive and responsive caregiving and secure attachment in the Strange Situation, although the effects observed in those replication studies were considerably weaker than those reported by Ainsworth and colleagues. Three separate meta-analytic studies summarizing this body of work all concluded that the replication studies, taken together, showed a moderate linkage between maternal sensitive responding and infant security. Therefore, it appears that maternal sensitive and

responsive caregiving plays an important, but not an exclusive, role in the formation of infant attachment security as assessed by the Strange Situation.

In addition to attachment security, the antecedents of attachment disorganization have been the subject of much theorizing and research. Hesse and Main posited that disorganized attachment is the result of frightening or frightened behavior on the part of the caregiver. The underlying assumption is that these circumstances present a very difficult situation for the infant because the same person who is supposed to help alleviate the child's fear is also the person who causes it or is also afraid and unable to help the child. This leads to an approach–avoidance conflict and to behavioral disorganization. Support for this proposition is found in studies of children who are maltreated. These children show significantly higher rates of disorganized attachment than do children who were not abused or neglected.

4.1.2. Parental Attachment Representations

Another variable that has been systematically linked to infant attachment security in the Strange Situation is maternal attachment security as assessed by the AAI. Thus, secure autonomous mothers typically have secure infants, dismissing mothers tend to have avoidant babies, and preoccupied mothers often have resistant infants. Finally, maternal unresolved loss or trauma as assessed by the AAI has been found to predict disorganized/disoriented infant attachment. This correspondence in attachment patterns, referred to as the "intergenerational transmission of attachment," has been found across many studies, even when the AAI was assessed prior to the infants' births. One study involving grandmothers as well as mothers and infants also found transmission of attachment across three generations. The correspondence between parental and child attachment is greater for mothers than for fathers, possibly because mothers typically serve as the primary caregivers.

How is attachment security passed on from parent to child? The main process suggested by attachment theory is through the quality of caregiving provided by the parent. Indeed, research has shown that a substantial part of the correspondence between parent AAI and infant Strange Situation classification (approximately 23–25%) is due to the level of maternal sensitive responsive caregiving. However, a large portion (75–77%) of the correspondence is not explained by quality of caregiving, at least as assessed in the relevant studies

(this discrepancy has been labeled the "transmission gap"). Thus, additional processes must play a role in the transmission of attachment security from generation to generation. Some of these processes may involve genetic transmission of attachment or child influences on attachment security.

4.2. Genetic Transmission of Attachment

Beyond caregiver sensitive responding, another mechanism that might underlie the transmission of attachment from parent to child is through their shared genes. Several twin studies have explored the potential genetic component of attachment development. Overall, these studies show a negligible role for genes, and a paramount role for the environment, in the formation of attachment security during infancy. However, it is possible that genes might be implicated in attachment security beyond infancy; one twin study with preschoolers found evidence for a small genetic effect. With respect to attachment disorganization, the pattern of findings has been inconsistent. Some studies (including a small-scale molecular genetics study) found evidence for heritability, whereas a recent twin study found no evidence for genetic influence on attachment disorganization. Further research is needed to elucidate the role played by nature on the formation and maintenance of attachments throughout the life cycle.

4.3. Child Influences on Attachment Security

4.3.1. Child Temperament

Although attachment researchers have traditionally focused on caregiver contributions to attachment security, infants themselves also play an important role in the formation of parent–child interaction patterns. Recall that attachment to caregivers is established during the first year of life, but long before then infants already exhibit their individuality in their affect and behavior, referred to as temperament. Many different aspects and components of early temperament have been identified, and their effects on the nature of parent–child interactions have been examined. Some aspects of temperament seem particularly relevant to infants' expression and regulation of attachment behavior, for example, proneness to distress (e.g., difficultness, emotionality, negative reactivity) and ability

to regulate distress effectively. Therefore, the potential influence of these temperamental features on the development of attachment security has been the subject of much debate and investigation.

In the strongest form of critique, some temperament researchers have argued that the Strange Situation reflects infants' temperamental characteristics rather than their accumulated experience regarding caregivers' responsiveness to their expressed needs. However, this analysis seems too biased in the direction of temperament effects and overlooks some important evidence. It appears that attachment classifications cannot be reduced to temperamental characteristics.

Today, it is generally acknowledged that both parent and child characteristics play a role in the formation of a secure attachment. In general, small to moderate associations between aspects of temperament and attachment security have been found, particularly when the AQS, rather than the Strange Situation, was used to assess attachment security. These studies show that infants who are more prone to distress early in life (e.g., difficult, reactive) generally show less secure base behavior later on.

However, beyond ascertaining the relative contribution of parent and child effects, it seems that the greater challenge for researchers during the upcoming years is to uncover how the two classes of antecedents interact in the formation of attachment. It seems likely that infants' temperament affects parents' ability to provide sensitive and responsive caregiving. Keeping highly irritable infants soothed and content is more demanding than caring for "easy" babies. Thus, infants prone to distress may elicit a lower level of sensitive responding, compared with calmer infants, from the same parents. Moreover, aspects of the environment (e.g., SES, family constellation, culture, support systems) also exert their effects on the two partners and the dyadic process. For example, parents who are exposed to high levels of stress or have little social support may be less able to provide responsive caregiving, particularly when caring for highly irritable infants.

4.3.2. Attachment in Children with Special Needs

In addition to temperament, other infant characteristics can influence the development of attachment security. Of special interest to researchers is the potential effects of infants' special needs (e.g., a serious medical condition, premature birth, developmental delay) on the formation of attachment security. It has been speculated that infants' medical or developmental problems (and parental perceptions of these problems)

might disrupt parent–infant interactions and thus hamper the development of secure attachments. However, studies have generally found that this is not the case. A meta-analysis summarizing this work showed that the majority of medically compromised infants were securely attached to their mothers. Yet, some conditions can delay the time course of the formation of early attachments. For example, studies on infants with Down's syndrome generally show that the emergence of early attachment is slower in this group.

5. ATTACHMENT THROUGHOUT THE LIFE SPAN: STABILITY AND CHANGE

Attachment theory posits that early attachment relationships serve, to a large extent, as a model for later close relationships. According to Bowlby, based on the early history of interactions with a caregiver, an infant develops internal representations or implicit ideas of both self and the attachment figure (e.g., a view of the self as worthy of love and of the mother as available and responsive). These internal representations, termed "internal working models," guide the individual's processing of relevant information and thus affect how the person approaches new situations. It is through these internal working models that early attachments are thought to influence later development and relationships. Because internal working models are believed to operate largely outside the realm of consciousness, they are considered quite resistant to change. However, such change is possible as a result of significant life events and new relationships. Thus, attachment theory hypothesizes that (a) there should be some degree of stability in individuals' attachment representations or internal working models throughout the life span and (b) change in the nature of these internal representations is not random but rather linked to important life events (negative or positive) or to new meaningful relationships (dysfunctional or constructive).

Findings from a number of independent long-term longitudinal studies generally reflect this complex mixture of stability and change. In these studies, attachment security was assessed during infancy using the Strange Situation and again during late adolescence or early adulthood using the AAI. Some of these studies found considerable stability in individuals' security status, whereas others found no such continuity. Importantly, stability of attachment appears to be linked to the

characteristics of the samples involved. Thus, stability of attachment from infancy to adulthood is more likely to be found in individuals brought up in stable and low-stressed families than for those raised in highly stressed and unstable family environments. Moreover, across various samples, change from early security to later insecurity was not arbitrary. Rather, such change was often linked to negative life events that likely affected the quality of caregiving experienced by the child (e.g., loss of a parent, divorce, serious physical or mental illness).

6. ATTACHMENT AND SUBSEQUENT CHILD OUTCOMES

With the gradual unfolding of attachment theory, and particularly since the Strange Situation was established as a valid measure of attachment during infancy, researchers began to examine the implications of early attachment relationships for later child outcomes. The underlying assumption in the majority of these studies is that positive attachment relationships (usually classified as secure) would be associated with positive outcomes and that negative attachment relationships (usually classified as insecure) would be associated with less positive or even negative outcomes. In other words, researchers hypothesized, and findings generally support the notion, that secure infants would subsequently show better adaptive skills in the areas of cognitive, emotional, and social development, whereas insecure infants' development may be associated with maladaptive characteristics in various domains.

6.1. Emotion Regulation

Emotional regulation refers to an individual's ability to manage and cope successfully with various levels of emotional arousal, including negative emotions, in response to external events and internal stimulation (e.g., anxiety, excitement). Current views suggest that the attachment system itself provides a basic mechanism for emotional regulation. Studies show that during the years when the attachment relationship is formed, children who develop secure attachment ties to their caregivers express a variety of emotions freely, whereas children who form insecure–avoidant attachment ties tend to inhibit emotions. Overall, studies following children from infancy to early childhood have found that secure children are spontaneous in their expression of a variety of emotions, read the emotional cues of others better, and generally tend to be more emotionally positive. Insecure children, particularly those with an avoidant classification, tend to exhibit minimal emotional expressiveness overall, and particularly restrain the expression of negative emotions. Also, these children tend to misperceive others' emotional displays, and to exhibit inappropriate affect considering the emotional climate of the circumstances (e.g., show anger when expressing positive emotions would be more appropriate).

6.2. Prosocial Behavior and Interpersonal Relationships

Longitudinal studies have found that a history of responsive care resulting in secure attachment is associated with a number of positive social developmental outcomes. Yet it is difficult to draw certain conclusions as to the role of early attachment security in later social competence because the parent–child relationship involves many aspects in addition to attachment. In general, findings suggest that children who were classified as securely attached during the first year of their lives were judged during the preschool years to be more socially competent and empathic toward their peers, as well as to have higher self-esteem, than children with a history of insecure attachment. During middle childhood, children who where classified as secure in their early attachment relationships were also found to be more socially accepted by their peers and more adept at forming close friendships than children who were classified as insecure. Moreover, secure infants grow to be adolescents who are more capable of creating intimacy than do insecure infants. However, the correlational nature of this evidence precludes any firm conclusions regarding the causal role played by attachment security. The early attachment relationship is only one component within a larger developmental context predicting later competence in interpersonal domains.

6.3. Cognitive Development

Studies have found relations between classification of attachment during infancy and specific cognitive functions during childhood. The most consistent evidence indicates that secure infants grow to have some cognitive advantages over insecure infants. Caution should be exercised in interpreting results regarding the relations between early attachment classification and later cognitive performance given that findings generally

refer to the social aspects of cognition (social cognition), particularly problem solving, reasoning, and information processing within social contexts. However, there is some evidence suggesting that secure infants grow to have better cognitive functioning in the areas of attention, memory, and generalization during childhood than do insecure infants.

6.4. Problem Behavior and Mental Health

Research has shown that children with an insecure attachment history are more likely to develop behavioral and emotional problems than are children who had secure attachment relationships during infancy. Children with an insecure–avoidant history were found to exhibit more negative "acting out" behaviors, as well as to suffer from more depressed moods and anxiety during the preschool and early school years, compared with children who were securely attached, although the magnitude of these effects was generally small. Moreover, an avoidant–disorganized attachment classification during infancy was found to predict hostile and aggressive behavior during the school years. As with the other domains of development, no research provides evidence that insecure attachment is the primary factor, or even one of the more important factors, underlying any psychological disturbance.

It has also been argued that early insecure attachment contributes to psychopathology during adolescence and adulthood. Most notably, disorganized attachment during infancy has been linked to greater psychopathology in general during late adolescence, as well as to a specific type of psychological disturbance, namely dissociative symptoms. Infants identified as disorganized during the Strange Situation were found to report more dissociative states and experiences during late adolescence, reflecting disturbances of memory, identity, awareness, and cognition (e.g., lack of memory for significant past events). In addition, studies examining adult attachment have found a high proportion of insecure AAI classifications among psychiatric patients, although it might be that the psychiatric problems are affecting these individuals' responses during the AAI (rather than these individuals' state of mind with respect to attachment, or their attachment history, being the cause of the disturbance). A recent review suggested that dismissive adult attachment style is associated with acting out or externalizing disorders, whereas preoccupied attachment is associated with depression and anxiety or internalizing disorders. More research is still needed to

elucidate the role played by early attachment in psychological disturbance later in life.

7. APPLIED ASPECTS OF ATTACHMENT

Since attachment theory was first articulated by Bowlby during the 1950s and 1960s, it has had a substantial impact on how early parent–child relationships are viewed. This, in turn, has shaped the thinking and practices of clinicians and policymakers. These influences are presented in this final section.

7.1. Interventions with Infants and Young Children

Although Bowlby's ideas, when first introduced, were met with strong resistance on the part of the psychoanalytic community, over the years practitioners have come to incorporate many of these notions into their clinical work. Specifically, because early secure attachment is associated with positive child outcomes, interventions that focus on the enhancement of attachment security in populations at risk for developing insecure attachment have been developed. There are two main approaches that guide these intervention attempts. The first aims at improving the quality of parental caregiving behavior, particularly parental sensitivity. The second approach is typically more long term and focuses on altering the parent's insecure mental representations or implicit ideas of attachment into more adaptive secure ones. Of course, the two approaches are not mutually exclusive, and both can be incorporated into the same intervention.

7.1.1. Enhancing Parental Sensitivity

Improvement of the quality of caregiving behavior is typically done by first observing how the mother interacts with her infant. Many programs, as well as individual clinicians, that practice mother–infant psychotherapy use videotaping of mother–infant interactions to learn through observation about the strengths and weaknesses of each mother's caregiving style. Portions of the videotaped material are then shown to the parent to identify moments where she interpreted her child's signals correctly and responded appropriately as well as moments where she ignored or misinterpreted her child's cues. More general patterns of parental responding to the infant are also identified and discussed. Through this process,

the clinician improves the parent's ability to read the child's various cues and signals and to respond more sensitively and appropriately to them. Examples of the use of this process in interventions are the STEEP (Steps Toward Effective Enjoyable Parenting) program developed at the University of Minnesota, the Infant Parent Program at the University of Michigan, and a preventive intervention study conducted in The Netherlands by van den Boom. In addition, similar principles of enhancing responsive caregiving through the provision of information, feedback, modeling, and support often guide individual clinicians when working with mother–infant dyads.

7.1.2. Changing Internal Working Models

Interventions and therapies that attempt to modify insecure internal working models of attachment typically involve parents discussing and reflecting on their experiences with their own early attachment figures (i.e., their parents). The goal is to gain insight into the effects of these early (often negative) childhood experiences on the parents' current caregiving practices. The underlying assumption is that working through these early anxiety-provoking experiences will lead to the adoption of more secure internal working models of attachment, thereby preventing the replication of the same maladaptive interaction patterns with their own children. Note that although this approach focuses on the parents' internal representational level, parenting behavior is also addressed, but without the intense behavioral focus involved in the first approach to intervention discussed in the previous subsection.

Overall, a meta-analytic review of intervention programs using attachment theory in individual, dyadic, or group settings suggested that short-term, behaviorally based interventions with a specific focus were more effective in improving mother–infant quality of interaction and attachment security than were interventions focused on parental internal representations. However, the magnitude of the effects was generally small. Moreover, enhancing maternal sensitivity was found to be a more consistent outcome of interventions than was the effect on infant attachment classification.

7.2. Policy

Attachment theory has had an impact not only on the work of researchers and individual clinical practitioners but also on policymakers. Early on in his career, Bowlby and his colleague, John Robertson, observed and documented the emotional difficulties that hospitalized children exhibited when forced to separate from their parents. At the time, parents were not allowed to stay with their sick children, and children where forced to spend days and weeks with only brief, highly controlled visits from their parents. A filmed documentation of one hospitalized child's grief and pain over her separation from her parents had a tremendous impact on hospitals' practices and policies. Today, in most hospitals, parents are allowed and encouraged to spend unlimited time with their hospitalized children.

Child care policies were also largely affected by Bowlby and his followers' views on the importance of high-quality caregiving. One important area is the care provided for orphans and children who need to be removed from their parents' care. In the past, these children were typically put in institutions such as residential nurseries and orphanages. This practice was criticized by attachment theorists because the emotionally detached style that characterized care in these institutions did not enable the children to form meaningful relationships with consistent caregivers. As a response to the criticism of institutionalization, there has been a growing tendency to rely on foster care in an attempt to promote more intimate personalized caregiving within a family context. However, this alternative often appears to be less successful than expected because many children are passed on from one foster family to another for short amounts of time.

Finally, another area in which attachment theory has had an influence is how nonmaternal care during the day, and particularly group day care for infants and young children, is perceived and evaluated. In general, research shows that nonmaternal care, including group day care, does not, in and of itself, disrupt children's attachment security to their parents. However, low-quality day care—marked by large group size, high caregiver–child ratios, high turnover of caregivers, and lack of personalized caregiving—can be detrimental to infant–mother attachment security, particularly when other risk factors are present. This highlights the importance of providing infants with good quality day care settings to enable them to form meaningful relationships with consistent caregivers.

See Also the Following Articles

Child Development and Culture ■ Emotion ■ Prosocial Behavior, Development of

Further Reading

Atkinson, L., & Zucker, K. (Eds.). (1997). *Attachment and psychopathology*. New York: Guilford Press.

Bowlby, J. (1988). *A secure base*. New York: Basic Books.

Bretherton, I., & Waters, E. (Eds.). (1985). *Growing points of attachment theory and research. Monographs of the Society for Research in Child Development*, *50*(1–2, Serial No. 209).

Cassidy, J., & Shaver, P. R. (Eds.). (1999). *Handbook of attachment: Theory, research, and clinical applications*. New York: Guilford.

Goldberg, S. (2000). *Attachment and development*. London: Arnold.

Goldberg, S., Muir, R., & Kerr, J. (Eds.). (1995). *Attachment theory: Social, developmental, and clinical perspectives*. Hillsdale, NJ: Analytic Press.

Greenberg, M. T., Cicchetti, D., & Cummings, E. M. (Eds.). (1990). *Attachment in the preschool years*. Chicago: University of Chicago Press.

Parks, C. M., Stevenson-Hinde, J., & Marris, P. (Eds.). (1991). *Attachment across the life cycle*. London: Routledge.

Attention

Amir Raz

Columbia University and New York State Psychiatric Institute, New York, USA

GLOSSARY

alerting Achieving and maintaining a state of high sensitivity to incoming stimuli.

attention The mental ability to select stimuli, responses, memories, and thoughts that are behaviorally relevant among a host of others that are behaviorally irrelevant.

attentional networks Neural circuits subserving attentional processing that preserve a degree of anatomical and functional independence but interact in many practical situations.

cognitive control Processes such as conflict resolution, error correction, inhibitory control, planning, and resource allocation.

executive The mechanisms for monitoring and resolving conflict among thoughts, feelings, and responses (e.g., an attentional system concerned with such tasks as working memory, planning, switching, and inhibitory control).

neuroimaging Technological advances, often noninvasive, that permit tapping neurophysiological aspects of the behaving brain.

orienting The selection of information from sensory input.

top-down modulation A downstream (vs bottom-up) effect (e.g., cognitive control).

self-regulation Key mediator among genetic predisposition, early experience, and adult functioning (e.g., in controlling the reaction to stress, the capacity to maintain focused attention or the ability to interpret mental states both internally and in others).

Attention, endowed with a rich set of empirical findings, is a central theme in psychological research. Since the 1980s, human neuroimaging studies have allowed examination of the whole brain during tasks involving attention and, consequently, have provided much information on how the brain houses attentional processes. Brain networks subserving attention have been identified and described both anatomically and functionally. It is now possible to use these networks as model systems for the exploration of symptoms arising from various forms of pathology. This article outlines selective attention, its phenomenology, and its constituent attentional networks.

1. HISTORICAL CONTEXT

Attention, one of the oldest and most central issues in psychological science, is the problem of selecting for active processing certain aspects of the physical environment (e.g., objects) or ideas in a person's mind (which are stored in his or her memory). Many great minds have wrestled with the definition of attention. In 1890, William James wrote, "Everyone knows what attention

is. It is the taking possession of the mind in clear and vivid form of one out of what seem [to be] several simultaneous objects or trains of thought." James's account heavily joined attention with subjective experience. Moreover, James's effort to deal with both attention to objects and attention to "trains of thought" is important for understanding current approaches to sensory orienting and executive control. However, because attention in the sense of orienting to sensory objects can actually be involuntary and can occur unconsciously, attention is not, as the quote from James implies, the same as being aware.

Behavioral psychology postponed research into the internal mechanisms of attention after 1900. The quest for attentional mechanisms resumed following World War II, when Donald Broadbent proposed a filter, limited for the amount of information (in the formal sense of information theory), that was located between highly parallel sensory systems and a limited-capacity perceptual system. This view was of immense aid in making possible objective studies of limitations in the human ability to deal with more than one signal at a time in a variety of practical tasks.

As psychology moved toward the study of cognitive mechanisms, the new objective methods made it possible to ask about the processes of selection. It was found that words could activate their semantic associates even when there was no awareness of the words' identities (i.e., priming). The highly parallel organization found for sensory information extended to semantic processing. The act of selecting a word meaning for active attention appeared to suppress the availability of other meanings of the selected item and of competing items. More recently, it became increasingly clear that although attention can be contextualized as a form of "alertness" as well as an index to resource allocation/limitation, it has been viewed less as a filter or bottleneck and more as a mechanism for providing priority for motor acts, consciousness, and some kinds of memory.

Although the psychology of attention furnished a number of interesting results about the limits of performance and of unconscious processing, there was no agreement on whether attention involved separate mechanisms from those used to process data, let alone any analysis of what these mechanisms might be in neural terms. Advances in the understanding of neural systems underlying attention developed from experimental paradigms involving selection of sensory information and more recently from the technological innovation of noninvasive tools for imaging the living brain.

Brain imaging has forged an impressive link between psychology and neuroscience. As early as 1990, some scholars suggested that the human brain entertains several attentional systems of different but interrelated functions (e.g., orienting, target detection, alerting). More recently, it has been demonstrated that distinct brain areas indeed mediate different attentional processes and that it is possible to examine selective attention as an organ system with its own functional anatomy, circuitry, and cellular structure. Although still incomplete, this information has illuminated multiple important questions in cognitive science and has provided insights into neurological and psychiatric disorders of both children and adults. Attention allows humans to exercise voluntary control over thoughts, feelings, and actions. Variations in the operational efficiency of these attentional systems serve as a basis for differences in self-regulation and emotional control, and they promise to help describe mechanisms of volition and sustained effort.

2. GROSS CHARACTERISTICS OF ATTENTION

Humans are largely visual animals, perhaps due to a common adaptation to visual predation generally found in primates. Researchers and clinicians have investigated the optics, anatomy, development, pathology, and underlying neural processes of the visual system, making it the most widely studied perceptual system. Therefore, it is didactically advantageous to discuss attention within the visual realm. Visual attention allows humans to move attention around to various areas of the visual field and to change the detail with which they look at any given area. For example, a reader can look at this page and pay attention to its setup as a whole, or he or she can zoom in on specific words and certain letters therein. If the reader is paying attention to single characters, he or she can glean a lot of information about punctuation marks, catch typos, and even spot minute imperfections on the physical sheet. But in that case, the reader might miss the idea that a paragraph is trying to convey. A person has the ability to change the location of attention and also to change the size of the attentional focus. This has been called the "zoom lens" idea, or the "attentional spotlight," and relates to humans' common experience concerning the kind of attention needed for reading versus proofreading.

Given a large visual array of individual features, a person can have a choice of examining it globally or instead examining its specific features. The person can shift back and forth between them by changing the

focus. Some patients have difficulty in examining the local features; these patients usually have damage to the left temporo-parietal lobe. Other patients may do well with the local features but fail to get the overall contour; they usually have damage to the right temporo-parietal lobe. The parietal lobe tends to emphasize the shifting between local and global, whereas the temporal lobe seems to determine whether one can actually examine a local feature or a global feature of the stimulus.

People usually foveate, or look at, exactly the thing in which they are interested, and that generally relates people's attention to where they fixate. However, it is easy to dissociate the two: to cue people to attend to some location in space other than the center of gaze and then to show that they are very sensitive (i.e., have a low threshold or fast response time) to information that occurs at the cued location and are relatively slow or insensitive to information that occurs at the fovea. These covert shifts are believed to be used to select the part of the visual field to which people usually want to move their eyes. In everyday life, people usually follow covert shifts of attention with eye movements. Attention to visual elements can also apply to other modalities (e.g., auditory).

When multiple people talk simultaneously, it is sometimes necessary to select one of these streams of conversation to follow it in detail. An individual usually does that based on the location of the other person by visually orienting toward the person and/or honing in on the frequency of his or her voice (e.g., it easier to separate a male voice from a female voice than to separate two male voices) or on the content of the information (e.g., following by content). When an individual attends to one stream, the other information goes into the background; it is present but does not reach focal analysis.

There are data indicating that much of this unattended information is processed in subtle and complicated ways. Unattended information can suddenly get interesting because a person's name is mentioned or because something happens that is related to the conversation he or she is following and the person then finds himself or herself orienting to the new information. These phenomena have been studied experimentally in some detail.

3. ATTENTION AND PERCEPTION

The psychophysics literature provides good accounts of how visual thresholds correlate with attentional investment. However, improvement in "visual acuity" is not synonymous with altered thresholds for detection, better performance, or faster reaction times. For example, acuity requires the resolution of detail, whereas detection thresholds and reaction time can involve the summation of luminance, and this might obscure detail.

There is evidence that attention improves performance in spatial resolution tasks. Cognitive scientists draw a distinction between how attention may be useful for simple detection of events and how performance can improve at those events. Although performance may improve on increased attentional investment, there has been great controversy over what orienting attention to a sensory (e.g., visual) stimulus actually does. There is general agreement that the attended stimulus receives priority, so that reaction time to it is usually faster. For example, in the visual modality, there is evidence of enhancement of brain electrical activity over extrastriate visual areas by approximately 90 milliseconds after visual presentation. On the other hand, attention is not a panacea to perception, and there is a great deal that attention cannot do. For example, it is clear that attention to a peripheral stimulus does not compensate for the lack of acuity that would be present for a foveal stimulus. Stimuli in the fovea always have an advantage in detail, although the priority for processing the input has been placed elsewhere. Thus, while orienting to a location, attention can give priority to that location (i.e., targets that appear there will be perceived more rapidly and with lower thresholds), but it cannot substitute for the acuity provided by the fovea. Although the fovea is critical for acuity, the costs in reaction time for an unexpected foveal stimulus are just as great as those for an unexpected peripheral event. Thus, visual attention influences priority or processing preference.

Whereas investing attention is frequently associated with looking directly at the scene of interest, covert attention is the ability to select visual information at a cued location, without eye movements, and to grant such information priority in processing. Researchers have shown that the performance improvement at attended locations results, to some extent, from an enhanced spatial resolution at the cued location. Findings from further psychophysical studies support the hypothesis that attention increases resolution at the attended location. Studies exploring the relationship between visual attention and contrast sensitivity show that covert attention not only improves discriminability in a wide variety of visual tasks but also can speed up the rate at which information is processed. There are findings indicating that a person's contrast

sensitivity is greater in the lower visual meridian (not the higher one). The bulk of the evidence sets limits to the effects of attention on spatial resolution and specifies that certain visual (not attentional) constraints determine aspects of spatial resolution.

4. ATYPICAL ATTENTION

Attentional performance is affected by the biological rhythm. Diurnal reductions in attention normally occur during the hours of maximum sleepiness (2–7 AM) when body temperature is at a nadir, and enhanced performance is usually seen during the evening when body temperature peaks. During sleep, voluntary attention is often considered to be markedly attenuated or indeed absent. However, there is evidence to suggest that certain attentional, as well as preattentional, mechanisms remain intact. Dreaming is usually divorced from a sense of controlled awareness, but purported accounts of lucid dreaming (i.e., dreaming while knowing that one is dreaming) suggest that some control mechanisms may be available during sleep. Other common anecdotes include the incorporation of ambient sound into the dream content and the idea of sensitivity to a person's own name.

One of the ways in which to investigate information processing in sleep involves electrical recordings from the scalp (EEG). By averaging the brain's electrical response potentials to stimuli (ERP), it is possible to examine the processing capability of the sleeping brain. One such component to examine is the mismatch negativity (MMN). The MMN is an electrophysiological manifestation of involuntary preattentive processing in response to oddball stimuli. In a typical MMN paradigm, a "deviant" auditory stimulus is infrequently interspersed within a sequence of "standard" auditory stimuli. The MMN is evident in the difference waveform resulting from the subtraction of the ERP elicited by the standard stimulus from that elicited by the novel auditory stimuli (the deviants). The difference waveform, occurring even without attention, normally peaks between 100 and 250 milliseconds from the onset of the deviant event, depending on the dimension of deviance and its magnitude. The MMN is presumably associated with a mechanism that compares the current auditory input with the memory traces formed by previous auditory inputs and signals the occurrence of a mismatch.

In adults, MMN tends to decline during drowsiness. Whether it persists into adult human sleep is still debated, but other EEG components do reflect the brain's reaction to novelty. Although active midbrain inhibition blocks cortical activity in the developed brain, there is reason to believe that the sleeping infant brain is not as capable of blocking and inhibiting information efficiently. Indeed, MMN is obtainable from newborns and young infants, and there are nascent experimental data showing that the brain can learn (e.g., certain language contrasts), even while asleep, during those early developmental stages.

Another special cognitive state sometimes confused with sleep is hypnosis. Some people (e.g., highly hypnotizable individuals) may experience attentional and perceptual changes, which might not typically occur during common awareness, following particular suggestions. Similar phenomena can occur in certain patient populations. For example, within vision, hypnotic suggestions have been demonstrated to induce tunnel vision, color-blindness, hallucinations, alexia, and agnosia. Such phenomena can manifest in other modalities as well (e.g., auditory).

Hypnosis has been used clinically for hundreds of years and is primarily a phenomenon involving attentive receptive concentration. Clinicians practicing hypnosis suggest that when a person is in a hypnotic state, attentional and perceptual changes may occur that would not have occurred if the person had been in a more usual state of awareness. In a responsive individual, hypnotic perceptual alteration is accompanied by reproducible changes in brain action. For example, the activity of the anterior cingulate to painful stimuli or conflict resolution can be modulated by hypnotic suggestion. Most children are highly hypnotizable and are more easily inducted into the hypnotic state than are adults.

5. SUBSTRATES OF ATTENTIONAL NETWORKS

Attention can enhance neural processing at multiple levels. It is a selective aspect of information processing; some things are privileged and some things are ignored. The modulation of neural response by attention has been studied vigorously. For example, one recent hypothesis suggests that as processing becomes more complex, it becomes advantageous to spread processing across more neural areas (e.g., both hemispheres). Another experimental approach carefully crafted a behavioral task to probe and assess at least three distinct attentional networks: (a) achieving and maintaining the alert state, (b) orienting to sensory objects, and (c) selecting among conflicting actions. Behavioral and imaging data suggest that these networks

are independent. It is further possible to identify the neuroanatomy subserving these attentional networks.

Alerting involves a change in the internal state in preparation for perceiving a stimulus. The alert state is critical for optimal performance in tasks involving higher cognitive functions. Neuroimaging studies have shown activity in the frontal and parietal regions, particularly of the right hemisphere, when people are required to achieve and maintain the alert state even for a brief period. Lesions of these areas will reduce the ability to maintain alertness. Right frontal lesions have been shown to impair ability to voluntarily sustain attention, producing a larger number of errors over time than is found for left frontal patients in tasks involving continuous performance. Right parietal patients show deficits in maintaining the alert state and difficulty in attentional orienting that together produce a profound neglect in the visual field opposite the lesion. Alerting is thought to involve the cortical distribution of the brain's norepinephrine system arising in the locus coeruleus of the midbrain.

The orienting network involves the selection of information from sensory input. Cholinergic systems arising in the basal forebrain play an important role in orienting. The pulvinar, superior colliculus, superior parietal lobe, and frontal eye fields are often activated in studies of the orienting network. Orienting can be reflexive (e.g., when a sudden target event directs attention to its location), or it can be more voluntary (e.g., when a person searches the visual field looking for some target). Orienting typically involves head and/or eye movements toward the target, that is, overt orienting. However, it can also be covert. A few dorsal brain areas, including the superior parietal lobe and temporo-parietal junction, serve as a common source of attention to sensory stimuli. They produce effects within a network of areas that depend on modality (e.g., ventral visual areas in the case of visual input). The strongest evidence for localization of mental operations stems from the area of attentional orienting toward sensory stimuli, where a confluence of methods and experimental sophistication has demonstrated how separate brain areas can be invoked to organize a simple attentional shift. There is agreement that orienting of attention to a visual stimulus produces amplification in prestriate regions and that this affects processing in all subsequent regions and feeds back to influence processing in the primary visual cortex and perhaps in the lateral geniculate nucleus of the thalamus.

Executive control of attention involves more complex mental operations in monitoring and resolving conflict among computations occurring in different brain areas. Executive control is most needed in situations that involve planning or decision making, error detection, novel (or not well-learned) responses, and conditions judged to be difficult or dangerous as well as in overcoming habitual actions. The anterior cingulate and lateral frontal cortex are target areas of the ventral tegmental dopamine system. Brain imaging data have shown repeatedly that the anterior cingulate cortex (ACC) is an important node in this network. Toward that end, a number of neuroimaging studies have shown activation of the dorsal anterior cingulate in tasks requiring people to respond either to a prepotent response or to a rather strong conflicting dimension. For example, in the classic Stroop task, experienced readers name the ink color of a displayed word. Readers are usually slower and less accurate in responding to the ink color of an incompatible color word (e.g., the word "red" displayed in blue ink) than in identifying the ink color of a control item (e.g., the word "lot" displayed in red ink). Another task involves participants responding to the direction of a central arrow when flanking arrows could point in either the same (congruent) direction or the opposite (incongruent) direction. The Attention Network Test uses this flanker task to measure conflict and shows strong activation in the dorsal ACC. Neuroimaging studies have shown that these conflict tasks activate midline frontal areas (e.g., ACC), lateral prefrontal cortex, and basal ganglia. These experimental tasks provide a means of fractionating the functional contributions of areas within the executive attention network. Chiefly, the ACC was more active on incongruent trials than on congruent trials. This result could reflect the general finding that lateral areas are involved in representing specific information over time (i.e., working memory), whereas medial areas are more related to the detection of conflict.

Patients with focal brain lesions of the ACC initially display deficits of voluntary behavior. The notion of ACC involvement in cognitive control and volition has been a topic of much interest recently. Based on behavioral, optical, and neuroimaging data, it was reported that effective suggestion (i.e., verbal exhortation) prevented word reading and modulated focal brain activity in highly suggestible individuals. This top-down influence was both potent and selective in that it annihilated the Stroop interference effect and reduced ACC activity, respectively. Interpretation of these data implies that, at least in highly suggestible individuals, attentional manipulations can influence aspects of self-regulation by affecting neural activity in specific brain areas. Although attention and self-regulation have arisen

within two different research traditions, there have been some recent efforts to integrate these two directions considering hypotheses about specific neural mechanisms involved in both attention and cognitive control. For example, these prospective approaches advocate a synthesis that is likely to be productive in linking brain systems to aspects of child socialization.

See Also the Following Articles

Environmental Versus Individual Risk Taking: Perception, Decision, Behavior ■ Interpersonal Perception

Further Reading

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of atten-tional networks. *Journal of Cognitive Neuroscience*, *14*, 340–347.

Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. Ann Neurol, 10(4), 309–325.

Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of cognitive operations in the human brain. *Science*, *240*, 1627–1631.

Posner, M. I., & Raichle, M. E. (1994). *Images of mind.* New York: Scientific American Books.

Posner, M. I., & Rothbart, M. K. (1998). Attention, self-regulation, and consciousness. *Philosophical Transactions of the Royal Society of London B*, *353*, 1915–1927.

Raz, A., Shapiro, T., Fan, J., & Posner, M. I. (2002). Hypnotic suggestion and the modulation of Stroop interference. *Archives of General Psychiatry*, *59*, 1155–1161.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolu-tion. *Nature*, *396*, 72–75.

Attention and Concentration Training in Sport

Aidan Moran

University College, Dublin, Ireland

GLOSSARY

attention People's ability to focus on information derived either from the external world or from internal sources such as memory and imagination.

concentration The ability to focus effectively on the task at hand while ignoring distractions.

divided attention The ability to perform two or more concurrent actions equally well.

mental practice The systematic use of mental imagery to rehearse actions covertly before they are executed physically.

meta-attention People's informal theories about how their attentional systems work.

performance goals Specific tasks or actions that lie within the control of the athlete.

preperformance routines Sequences of preparatory actions that athletes tend to engage in before they execute self-paced skills.

selective attention The ability to "zoom in" on task-relevant information while blocking out distractions.

simulation training The use of simulated sport situations in practice in an effort to help athletes become accustomed to anticipated pressures and distractions.

trigger words Instructional cues that remind athletes what to focus on in competitive situations.

The term "attention" refers to people's ability to focus on information derived either from the external world or from internal sources such as memory and imagination. In sport psychology, attentional processes such as "concentration," or the ability to focus mental effort on the task at hand while ignoring distractions, are regarded as vital determinants of successful athletic performance. Given this importance of attention in sport, a variety of psychological exercises and techniques have been postulated to enhance athletes' concentration skills. Although none of these interventions has been validated adequately so far, some theoretical support exists for the use of the following strategies: simulation training, performance goal setting, preperformance routines, trigger words, and mental practice. Following an explanation of each of these concentration interventions, this article considers some unresolved issues along with potentially fruitful new directions for research in this field.

1. "ATTENTION" AND "CONCENTRATION" IN SPORT

Cognitive sport psychology is concerned with the scientific study of how the mind works in athletic settings. Within this field, the term "attention" refers to people's ability to focus on information derived either from the external world or from internal sources such as memory and imagination. This ability is regarded as a multidimensional construct that has at least three key components that can be explained as follows. First, when an athlete is asked to "pay attention" to an instruction provided by a coach at a training session, the athlete is required to concentrate on, or to exert mental effort in absorbing, the information presented. The second component of attention involves selective perception and occurs when sport performers "zoom in" on task-relevant information while blocking out distractions. To illustrate, a soccer goalkeeper who is preparing to catch a corner kick must focus only on the flight of the ball while ignoring the potentially distracting movements of other players in the penalty area. Finally, attention involves "multitasking," that is, the coordination of simultaneous skilled actions. For example, when a basketball player dribbles with the ball while scanning the court for a teammate to pass the ball to, the player is engaging in a form of mental time sharing called "divided attention." In summary, the construct of attention refers to three distinct cognitive processes: concentration or effortful awareness, selectivity of perception, and the ability to coordinate two or more concurrent actions successfully.

Since the 1950s, two main theoretical models of selective attention have been postulated in cognitive psychology. First, the "spotlight" metaphor suggests that selective attention resembles a mental beam of light that illuminates targets located either in the external world around us (sensory information) or in the subjective domain of our own thoughts and feelings (mental events). Another popular theoretical approach is the "resource" model of divided attention. This metaphor suggests that attention is analogous to a pool of mental energy that is allocated to fulfill processing demands according to certain strategic principles. To illustrate, the principle of "automaticity" suggests that the more practiced a task is, the more automatic it becomes and the less attentional resources it requires. Based on this principle, it may be predicted that expert athletes are especially vulnerable to distractions because they have "spare" mental resources due to the automated

nature of many of their perceptual motor skills. The issue of how to refocus the minds of such performers is addressed later in this article.

Athletes, coaches, and sport psychologists regard attentional skills such as concentration, or the ability to focus effectively on the task at hand while ignoring distractions, as essential prerequisites of success in sport. This claim is supported by both descriptive and experimental evidence. First of all, surveys of expert sport performers such as U.S. Olympic athletes indicate that concentration is consistently rated among the most important mental skills required for sporting excellence. Next, qualitative studies suggest that total absorption in the task at hand is a key feature of "flow" or "peak performance" experiences in sport, that is, those coveted yet elusive moments during which athletes perform at the best of their abilities. Finally, there is growing experimental evidence that athletes (e.g., sprinters) who have been trained to use certain concentration strategies (e.g., deliberately focusing on task-relevant cues) tend to perform better than their counterparts who are assigned to nonfocusing control conditions. Taken together, these sources of evidence converge on the conclusion that the ability to pay attention to the task at hand is a vital determinant of success in sport.

Based on peak performance research, three key principles of effective concentration can be identified. First, a focused state of mind requires intentionality and a willingness to expend deliberate mental effort on the part of the athlete concerned. Put simply, one must prepare to concentrate effectively rather than merely hope that it will happen fortuitously. Second, although skilled athletes might be able to divide their attention efficiently between two or more concurrent actions, they can focus consciously on only one thought at a time. This principle stems from the fact that "working memory," or the cognitive system that regulates people's attentional deployment, is limited in its capacity and duration. Accordingly, it can be overloaded easily when several thoughts and/or behavioral intentions compete for conscious attention. Interestingly, this danger of cognitive overload is increased when athletes become preoccupied with winning a competitive encounter. In this situation, they may find themselves torn between trying to focus on performing an action and worrying about the likely outcome of the encounter in which they are participating. Finally, as mentioned earlier, research on flow states indicates that athletes' minds are focused optimally when they are totally absorbed in the task at hand. In such a state of

mind, there is no difference between what they are thinking about and what they are doing. In summary, athletes tend to concentrate most effectively when they prepare to exert mental effort and can aim their attentional spotlight at actions that are specific, relevant, and under their own control.

2. TRAINING ATTENTIONAL SKILLS IN ATHLETES: CONCENTRATION EXERCISES AND TECHNIQUES

Given the importance of attentional processes in sport, a wide range of strategies has been promulgated by applied sport psychologists in an effort to enhance athletes' focusing skills. In general, these interventions fall into two types of activities: concentration training exercises and concentration techniques. Theoretically, the main distinction between these activities is that the former are designed for use during athletes' training sessions, whereas the latter tend to be practiced most frequently in actual competitive situations. Not surprisingly, however, this distinction is blurred by the fact that athletes are increasingly aware of the value of replicating competitive situations in their training regimes. Accordingly, there may be an overlap between the concentration exercises and techniques they employ in training and in competition.

2.1. Concentration Exercises

One of the most popular focusing exercises recommended by sport psychologists is a visual search task called the "concentration grid." In this task, participants are presented with a random array of two-digit numbers (from 00 to 99) on a page and are required to scan and "tick off" as many digits as possible in a given sequence (e.g., beginning with 00 and increasing consecutively) within the specified time limit of 1 minute. The number of digits marked in the correct sequence within this duration serves as an alleged index of the athletes' concentration skills. Unfortunately, although the concentration grid may have some intuitive plausibility, it has not been validated as a concentration exercise either theoretically or empirically. Therefore, its conceptual rationale and practical utility remain unresolved.

In contrast to the grid, a more recent concentration exercise known as "simulation training" seems to have some merit. Briefly, this exercise is based on the proposition that athletes can learn to improve their concentration skills if they simulate real-life competitive situations in practice. Anecdotal testimonials to the value of this practice in various sports have emerged during recent years. For example, some leading soccer coaches have prepared their players for the pressure of penalty "shoot-outs" by requiring them to practice walking from the center circle of the pitch to the penalty area—as happens in actual match conditions. Unfortunately, despite its apparent face validity, simulation training has received little or no empirical scrutiny as a concentration strategy. Nonetheless, some support for its theoretical rationale may be found within the field of cognitive psychology. For example, research on the "encoding specificity" principle of learning shows that people's recall of information is facilitated by conditions that resemble those in which the original encoding occurred. Applying this principle to simulation training, it may be argued that replication of real-life conditions in practice inoculates athletes against anticipated pressures and distractions. By this reasoning, it is conceivable that simulation training could enhance athletes' concentration skills.

2.2. Concentration Techniques

The second category of concentration strategies recommended for use by athletes in competitive situations involves the use of psychological techniques such as (a) performance goal setting, (b) preperformance routines, (c) trigger words, and (d) mental practice. These techniques are explained in what follows.

2.2.1. Performance Goal Setting

In psychology, a "goal" is a target or an objective that people strive to attain such as sinking a putt, winning a match, or being selected for a given team. Therefore, "goal setting" is the process by which people establish desirable objectives for their actions. Typically, sport psychologists distinguish between "result goals" (e.g., the outcome of a competition) and "performance goals" (i.e., specific actions that are under the athlete's control). Using this distinction, some psychologists have speculated that the process of setting performance goals can improve athletes' concentration skills. For example, a golfer could improve his or her concentration on the course by focusing on specific "controllable" goals such as keeping his her head steady and/or maintaining a slow rhythmic swing. Empirical support for this conjecture springs from studies on the correlates of people's "best" and "worst" athletic performances.

In particular, research shows that collegiate athletes performed worst when they were preoccupied with result goals. Conversely, their best performances tended to occur when they adopted explicit performance goals. In summary, there is a legitimate theoretical rationale for the use of performance goals as a concentration technique by athletes.

2.2.2. Preperformance Routines

It has long been apparent that many athletes display distinctive idiosyncratic sequences of preparatory actions before they perform key skills, especially in individual sports. For example, top tennis players tend to bounce the ball a consistent and characteristic number of times before their first serve and a different yet equally consistent number of times before their second serve. These bounce patterns help the player to focus on the first step in his or her pre-serve preparatory sequence. Similarly, expert golfers tend to adhere to a preferred number of practice swings and "waggles" of the club before they strike the ball. These distinctive action sequences and/or repetitive behaviors are called "preperformance routines" and are especially prevalent prior to the execution of "self-paced" skills (i.e., actions that are carried out largely without interference from other people).

Three main types of routines are evident in sport. First, "pre-event" routines are general patterns of preparation that athletes display on the days preceding competitive action. Included here are preferences with regard to food and sleeping arrangements. Second, "pre-performance" routines are characteristic sequences of thoughts and actions that athletes adhere to prior to skill execution. Finally, "postmistake" routines are action sequences that may help performers to put errors behind them so that they can refocus on the task at hand. For example, a tennis player may "shadow" the correct movement of a volley that led to an error.

Support for the value of preperformance routines as concentration techniques comes from both theoretical and empirical sources. Theoretically, preperformance routines may improve concentration for at least two reasons. First, they may enable athletes to concentrate on the current moment rather than on past events or on possible future outcomes. In addition, they may prevent athletes from devoting too much attention to the mechanics of their well-learned skills. This habit, which is also known as "paralysis by analysis," can "unravel" automaticity and lead to a deterioration in performance. In other words, routines may help to

suppress the type of inappropriate conscious control that often occurs in pressure situations. Empirically, there is some evidence from case studies that routines can improve athletes' concentration skills and performance. For example, research suggests that more proficient golfers benefited more from using routines than did less skilled golfers.

Unfortunately, advocacy of preperformance routines gives rise to two practical issues that need to be addressed. First, routines may lead to superstitious rituals on the part of the performer. Thus, certain athletes feel compelled to wear "lucky" clothes (e.g., Tiger Woods likes to wear something red on the final day of a golf tournament) or to listen to "lucky" songs before important competitive events. Indeed, athletes' precompetitive behavior is often a mixture of routines and superstitions. A second problem with routines is that they need to be revised regularly as a precaution against the possibility that they will become automated through frequent practice. In other words, if athletes maintain the same preperformance behavior indefinitely, their minds may begin to wander as they proceed through the various steps of the routine. Because of this problem, applied sport psychologists face the challenge of helping athletes to develop consistent but not automated preparatory actions.

2.2.3. Trigger Words

Most athletes talk to themselves covertly as they train or compete. For example, gymnasts may use words such as "forward" to remind themselves to push their bodies upward as they attempt to master a floor routine. Similarly, tennis players may try to trigger a smooth service action by telling themselves to "reach and hit." In general, this internal speech (or "self-talk") may involve praise (e.g., "well done—that's good"), criticism ("you idiot—that's a stupid mistake"), or instruction ("swing slowly").

Typically, the purpose of instructional verbal triggers is to help athletes refocus their minds on task-relevant cues just before skill execution. For example, when standing over a putt, golfers may use phrases such as "steady head" to prevent themselves from looking up to see whether or not the ball has gone into the hole.

Theoretically, self-talk could enhance concentration skills by reminding athletes about the most important cues in any given situation. Interestingly, this strategy was used by Serena Williams during the 2002 Wimbledon women's singles tennis final in which she defeated her sister Venus. During this match, Serena

read notes to herself during the "change-over" time between games. Afterward, she explained that these notes had contained trigger words to remind her to "hit in front of you" or to "stay low."

Despite such anecdotal testimonials to the value of trigger words, few studies have tested the validity of self-talk as a concentration technique. However, it is generally believed that for optimal efficacy, instructional self-statements must be short, vivid, and positively phrased.

2.2.4. Mental Practice

The term "mental practice" or "visualization" refers to the systematic use of mental imagery to rehearse physical actions. Put simply, it involves "seeing" and "feeling" a skill in one's imagination prior to its actual execution. This strategy is used by golfers such as Tiger Woods when preparing to play a shot. Although there is considerable experimental evidence that mental practice facilitates skill learning and performance, few studies have been conducted on imagery as a concentration technique. Anecdotally, however, many athletes report using imagery to prepare for anticipated scenarios, thereby reducing the likelihood of being distracted by unexpected events.

3. UNRESOLVED ISSUES IN TRAINING ATTENTIONAL SKILLS IN SPORT

At least five issues need to be addressed before attentional and/or concentration training techniques can be used effectively with athletes in applied settings. First, before they attempt to train concentration skills, sport psychologists should try to find out why athletes lose their focus in the first place. Unfortunately, until recently, little research was available on the causes of distractibility in sport performers, especially distractibility that arises from "internal" sources such as athletes' own thoughts and feelings. However, some recent theoretical models have examined the role of unconscious factors in this regard. Augmenting this research, self-report scales have been developed by sport psychologists to assess athletes' susceptibility to self-generated cognitive interference. Second, sport psychologists have much to learn about the attentional demands of various sports. For example, whereas some activities (e.g., weightlifting) may require short periods of intense concentration, others (e.g., cycling) appear to demand

sustained alertness for longer periods of time. Intuitively, it seems unreasonable to expect that the same toolbox of concentration interventions will work equally well in all sports. Third, some researchers have raised the question of whether or not sport performers know exactly what they should be concentrating on in various athletic situations. For example, should a tennis player focus only on the ball during a rally, or should the player use cues from his or her opponent's behavior (e.g., position of feet) in an effort to anticipate the likely target of a subsequent shot? Unfortunately, this question has been largely neglected by sport psychologists in their enthusiasm to provide practical mental skills training programs for athletes. Fourth, what is the best way in which to measure concentration skills in athletes? Until this question has been answered empirically, it is difficult to evaluate whether or not concentration skills interventions are effective. Finally, no explicit criteria currently exist for evaluating the maintenance of attentional skills improvements over time. In the absence of such criteria, researchers have not yet ruled out the possibility that athletes' concentration patterns may return to preintervention levels despite athletes' participation in mental skills training programs.

4. NEW DIRECTIONS IN RESEARCH ON CONCENTRATION SKILLS TRAINING

At least four potentially fruitful new directions in research on concentration skills training in athletes may be identified. First, further research is required on the "meta-attentional" processes of athletes, that is, their theories on how their own concentration systems work. It may be argued that the entire program of concentration skills training in sport psychology is really an exercise in meta-attentional training, whereby athletes are empowered psychologically by gaining an understanding of, and control over, their own concentration processes. As yet, however, little is known about the nature, accuracy, and/or modifiability of athletes' meta-cognitive theories. Second, additional research is required on the relationship between the structure of various athletic activities and their attentional demands. For example, do untimed activities such as golf place different cognitive demands on athletes' concentration systems compared with those imposed by timed competitions such as soccer? If so, what theoretical mechanisms could account for such differences? Third, more research is needed to establish

how emotional factors (e.g., anxiety) affect athletes' concentration processes. This task could be achieved by exploring the effects of anxiety on the visual search behavior of athletes tackling laboratory simulations of sport-relevant tasks. Finally, research is required to evaluate the efficacy of simulation training and the various concentration techniques described earlier in improving athletes' focusing skills in competitive sport situations.

See Also the Following Articles

Attention ■ Goal Setting and Achievement Motivation in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Psychological Skills Training in Sport

Further Reading

Abernethy, B. (2001). Attention. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 53–85). New York: Macmillan.

Boutcher, S. H. (2002). Attentional processes and sport performance. In T. Horn (Ed.), *Advances in sport psychology* (2nd ed., pp. 441–457). Morgantown, WV: Fitness Information Technology.

Fernandez-Duque, D., & Johnson, M. L. (1999). Attention metaphors: How metaphors guide the cognitive psychology of attention. *Cognitive Science*, 23, 83–116.

Hatzigeorgiadis, A., & Biddle, S. J. H. (2000). Assessing cognitive interference in sport: Development of the Thought Occurrence Questionnaire for Sport. *Anxiety, Stress, and Coping*, 13, 65–86.

Jackson, R. C., & Baker, J. S. (2001). Routines, rituals, and rugby: Case study of a world class goal kicker. *The Sport Psychologist*, 15, 48–65.

Moran, A. P. (1996). *The psychology of concentration in sport performers: A cognitive analysis*. Hove, UK: Psychology Press/Taylor & Francis.

Moran, A. P. (2000). Improving sporting abilities: Training concentration skills. In J. Hartley, & A. Branthwaite (Eds.), *The applied psychologist* (2nd ed., pp. 92–110). Buckingham, UK: Open University Press.

Moran, A. P. (2003). The state of concentration skills training in applied sport psychology. In I. Greenlees, & A. P. Moran (Eds.), *Concentration skills training in sport* (pp. 7–19). Leicester, UK: British Psychological Society.

Moran, A. P. (2004). *Sport and exercise psychology: A critical introduction*. London: Psychology Press/Routledge.

Schmid, A., & Peper, E. (1998). Strategies for training concentration. In J. M. Williams (Ed.), *Applied sport psychology: Personal growth to peak performance* (3rd ed., pp. 316–328). Mountain View, CA: Mayfield.

Simons, J. (1999). Concentration. In M. A. Thompson, R. A. Vernacchia, & W. E. Moore (Eds.), *Case studies in applied sport psychology: An educational approach* (pp. 89–114). Dubuque, IA: Kendall/Hunt.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, 101, 34–52.

Williams, J. M., & Leffingwell, T. R. (2002). Cognitive strategies in sport and exercise psychology. In J. Van Raalte, & B. W. Brewer (Eds.), *Exploring sport and exercise psychology* (2nd ed., pp. 75–98). Washington, DC: American Psychological Association.

Attention Deficit Disorders: School-Based Interventions

George J. DuPaul and Katy E. Tresco

Lehigh University, Bethlehem, Pennsylvania, USA

GLOSSARY

antecedent-based intervention Any treatment strategy that involves manipulating environmental events prior to a target behavior occurring so as to increase or reduce the probability that the target behavior will occur; these interventions are considered to be proactive preventive strategies.

consequent-based intervention Any treatment strategy that involves manipulating environmental events after a target behavior occurs so as to increase or reduce the probability that the target behavior will occur; these interventions are considered to be reactive strategies.

curriculum-based measurement (CBM) Assessment of a student's math or reading skills using brief (2- to 3-minute) probes derived directly from the curriculum being taught in the classroom; CBM provides a cost-efficient and sensitive method of documenting a child's instructional level and changes in skills over time.

functional assessment Collection of interview and/or observation data to establish the environmental events that increase the probability that a target behavior will occur.

point of performance The exact time that a target behavior is most likely to occur and the setting where a behavior is likely to be exhibited.

Attention deficit/hyperactivity disorder (AD/HD) is associated with academic and social impairments that significantly compromise school performance. Effective school-based interventions include antecedent-based and consequent-based strategies that are implemented by teachers, peers, parents, computers, and/or the target students themselves.

1. DESCRIPTION AND DEFINITION OF ATTENTION DEFICIT/ HYPERACTIVITY DISORDER

Attention deficit/hyperactivity disorder (AD/HD) is a psychiatric disorder characterized by developmentally inappropriate levels of inattention, impulsivity, and/or overactivity. There are three subtypes of AD/HD: predominantly inattentive, predominantly hyperactive–impulsive, and combined type, with the latter being most common among clinic referrals and research participants. AD/HD affects approximately 3 to 5% of the school-aged population in the United States, with boys outnumbering girls by two to one in epidemiological

samples and by six to one in clinic samples. The symptoms of AD/HD frequently are associated with impairment in academic, social, and behavioral functioning. The disorder tends to be chronic and often is a precursor to the development of other disruptive behavior disorders, including oppositional defiant disorder and conduct disorder. Among the greatest long-term risks associated with AD/HD are scholastic underachievement, grade retention, and school dropout. The most effective treatments for this disorder include psychotropic medication, especially psychostimulants (e.g., methylphenidate), behavioral strategies applied at home and school, and (when necessary) academic interventions. It is assumed that most children with AD/HD will require a combination of these treatments to successfully ameliorate symptomatic behavior over the long term.

The most widely studied and used medications for AD/HD are central nervous system (CNS) stimulants such as methylphenidate (Ritalin, Concerta, and Metadate), combined stimulant compounds (Adderall), and dextroamphetamine (Dexedrine). CNS stimulants are associated with improved attention and impulse control in the majority of treated children with AD/HD. Side effects are benign in most cases, with insomnia and appetite reduction being the most common adverse effects. Behavioral effects vary considerably across individuals as well as across doses within individuals. Furthermore, stimulant medication effects are observed primarily during the school day because their duration of action ranges from 4 to 8 hours. For these reasons, it is important that information regarding behavior changes at home and school, as well as possible improvement or deterioration in academic performance, is taken into account when making medication decisions. For those individuals who do not respond to a CNS stimulant, alternative medications that may improve AD/HD symptoms include antidepressants (e.g., bupropion), antihypertensives (e.g., clonidine), and selective norepinephrine reuptake inhibitors (e.g., atomoxetine).

2. CONCEPTUAL FOUNDATIONS OF SCHOOL-BASED INTERVENTIONS

There are several principles that are critical to the design of effective school-based interventions for students with AD/HD: gathering assessment data that directly inform intervention planning, implementing interventions at the point of performance, individualizing intervention strategies for each student, using a balanced treatment plan composed of both antecedent-based and consequent-based interventions, using strategies to address both academic and behavioral difficulties, and employing multiple individuals to implement treatment components.

2.1. Linking Assessment Data to Intervention

The assessment of children suspected of having AD/HD not only should be directed at establishing a diagnosis but also should provide data that can directly inform treatment planning. Thus, although norm-referenced measures such as behavior rating scales are helpful for making diagnostic decisions, additional measures should be included to provide information about possible directions for treatment. Specifically, functional assessment procedures, such as observations of target behaviors (e.g., calling out in class without permission) in the context of environmental events (e.g., classmates laughing), can be used. Data from a functional assessment can help to identify antecedents and consequences that may be maintaining a target behavior and that can be manipulated to change the frequency of a behavior. For example, disruptive behavior that consistently leads to peer attention as a consequence can be reduced by designing an intervention that provides the student with peer attention for engaging in appropriate behavior (e.g., peer tutoring).

2.2. Intervening at the Point of Performance

Because children with AD/HD typically exhibit poor impulse control, their behavior is more likely to be controlled by immediate environmental events than by long-term contingencies. For this reason, the most effective intervention strategies will be those that are implemented as close to the "point of performance" as possible. For example, if an intervention is necessary to reduce disruptive behavior occurring in math class and the latter is conducted from 9:00 to 9:45 AM each school day, the most effective interventions will be those that are implemented in math class from 9:00 to 9:45 AM. Treatment procedures that are removed in time and place from the point of performance (e.g., weekly counseling sessions in the school psychologist's office) will be less effective.

2.3. Individualizing Intervention Strategies

Children with AD/HD exhibit a wide range of possible symptoms, disruptive behaviors, and academic difficulties. Thus, a "one size fits all" approach to treatment, wherein it is assumed that all students with AD/HD respond the same way to the same interventions, is likely to fail. The selection of intervention components should be made on the basis of individual differences in symptom severity, presence of comorbid conditions (e.g., oppositional defiant disorder), possible functions of the target behaviors, teacher acceptability of prescribed interventions, and empirical evidence supporting specific treatments for the target behaviors.

2.4. Combining Antecedent and Consequent Interventions

The application of both antecedent and consequent intervention strategies allows for students to have clear expectations of what is expected of them in advance (antecedent) as well as what will occur if problem behavior arises (consequent). Combining antecedent and consequent interventions also provides for more positive interactions and experiences, thereby potentially increasing the chances that the student will behave appropriately. Interventions based solely on aversive consequences or punishments often result in increased negative interactions between the student and the teacher and minimize the opportunity for positive reinforcement.

Antecedent or proactive intervention strategies are those that are implemented to prevent or preempt academic and/or behavioral difficulties from occurring. These could involve implementing classroom strategies and rule systems to provide structure, making environmental modifications (e.g., placing a child with academic or behavioral difficulties near the teacher or near children who are less likely to interact negatively with the child), having a child repeat instructions before beginning an assignment or activity, and/or providing limited task choices to a child prior to beginning seatwork.

Alternatively, consequent or reactive strategies are those implemented to respond to both appropriate and problem behavior when it does occur. An example of a consequent intervention might be a response cost system in which the student is "fined" for exhibiting problem behavior (e.g., loses points or reinforcement)

yet also has the ability to earn and receive positive feedback when appropriate behavior is exhibited.

2.5. Combining Behavioral and Academic Interventions

Children with AD/HD often exhibit both academic and behavioral difficulties in the classroom. It is important to address both areas because a student who is frustrated academically may be more likely to exhibit negative behavior. Based on the results of a 1997 meta-analysis performed by DuPaul and Eckert on studies of school-based interventions from 1971 to 1995, improving a child's academic skills may potentially improve his or her behavior in the classroom as well. In addition, targeting the attention difficulties associated with AD/HD and improving on-task behavior through behavioral intervention may potentially have a positive impact on academic skills. By including strategies targeting both academic skill areas as well as relevant behavior problems, teachers may be able to maximize the effectiveness of the interventions implemented. Collecting curriculum-based measurement and functional assessment data can lead to a better understanding of the factors contributing to academic and behavioral problems. Identifying the specific instructional needs and environmental variables that reliably proceed or follow inappropriate behavior can be a key component in the formulation of effective interventions

2.6. Using Multiple Intervention Agents

Effective intervention strategies can be implemented by a variety of mediating or intervention "agents." Interventions can be mediated by teachers, peers, computers, parents, and/or the students themselves. By using multiple agents, sole responsibility for addressing all academic and behavioral difficulties associated with AD/HD does not fall on the teacher. Instead, multiple resources can be used to provide a more comprehensive, and potentially more cost-effective, method to support teacher instruction as well as to deliver interventions. Table I provides some possible intervention strategies listed by mediating agent. Specific intervention strategies organized by mediating agent and setting (elementary school vs secondary school) are discussed in more detail in later sections.

Mediating agent	Elementary school	Secondary school
Teacher	Classroom rules	Contingency contract
	Response cost systems	Direct instruction in study/organizational skills
	Environmental manipulations	
	Instructional modifications	
	Instructional strategies	
Peer	Peer tutoring	Peer coaching
	Group contingencies	
Computer	Instruction	
	Drill and practice	
Parent	Goal setting	Home-based reinforcement
	Contingency contracting	Homework program
	Home-based reinforcement	
	Parent tutoring	
Self	Self-monitoring	Self-monitoring
		Self-evaluation/ Reinforcement

Source: From DuPaul, and Power (2000). Copyright 2000 by the American Psychiatric Association. Adapted by permission.

3. INTERVENTIONS FOR ELEMENTARY SCHOOL-AGED STUDENTS

As described previously, combinations of interventions appear to be most effective in improving and maintaining appropriate classroom behavior. This section discusses intervention strategies for improving academic and behavior performance using both antecedent and consequent intervention strategies across mediating agents. Typically, the most effective interventions include the use of multiple strategies and intervention agents.

3.1. Teacher-Mediated Strategies

Classroom strategies that have previously been found to be effective in improving the academic and behavior performance of children with AD/HD include teaching classroom rules, using response cost systems, altering the structure of the classroom environment, adjusting academic expectations to match students' current skills or instructional level, modifying instructions, and arranging for choices in academic tasks.

Most classrooms have rules that are typically reviewed with students at the beginning of the academic school year. Periodic review and reteaching of these rules throughout the year can be an effective support system for maintaining appropriate behavior. Teachers should develop four or five positively phrased classroom rules. By stating them in a positive manner, the rules themselves help to inform students as to how to behave instead of providing an example of what "not" to do. Initial teaching should include both examples and nonexamples of following the rules, and examples should be elicited from the students to ensure their understanding. Periodic review should include a few minutes spent each week reviewing one or two of the rules and their importance and again eliciting examples from students to confirm their understanding. This frequent review and discussion of rules enables students to have a clear understanding of what is expected of them. In addition, pairing classroom rules with a response cost system can be an effective combination of antecedent- and consequent-based strategies.

A response cost system requires a clearly defined rule system that is reviewed on a frequent basis. Within a response cost system, students may earn "tokens" (e.g., points, chips) for following rules and can lose tokens when rules are broken. The tokens can then be applied toward privileges such as taking the attendance to the office, passing out papers, and purchasing items from a "class store." For a response cost system to be effective, students must have clear expectations of when the rules are in effect and what behavior will result in both the earning and losing of tokens. Examples of both situations should be given to confirm students' understanding while the rules are being reviewed. Students should also understand what privileges are available, when they may exchange their tokens for privileges, and how many tokens are necessary for each privilege. Student motivation is a key component of an effective response cost system. Therefore, it may be beneficial to have students participate in a discussion of which privileges will be available and at what cost. It may also be helpful to determine ahead of time what privileges students are working for and to remind them periodically.

It is important to note that although classwide systems are widely effective for most students in the classroom,

individual modifications may be necessary for those students whose behavior problems are more severe. If a student typically breaks 20 rules in a 30-minute class period, asking the student to break only 2 rules may be an unrealistic expectation and the child may be unmotivated by the system. Therefore, the amount of tokens necessary to earn a privilege might need to be altered on an individual basis. Similarly, if students can earn points for following rules for a certain amount of time, it might be necessary to decrease the required duration for specific students to allow them to earn points. Adjusting the time period also provides for more immediate feedback to those students who might need more frequent reinforcement.

In addition to classroom rules and response cost systems, children with AD/HD may benefit from alterations to the classroom environment to allow for increased monitoring of classroom behavior. Children with AD/HD may benefit from being placed near the teacher or with peers who are less likely to interact negatively with them. In addition, although it may seem advantageous to place children with severe attention and behavior difficulties in isolated areas of the classroom, it is important that they remain a part of important classroom activities and have access to teacher instruction and positive peer interactions.

Behavior and academic difficulties often associated with AD/HD may be exacerbated by instructional demands that are too high for students' current level of functioning. Frustration with tasks that are too difficult may increase the likelihood that some children will exhibit negative behaviors. Thus, it is important to assess and adjust academic expectations to match a student's current skills or instructional level. Curriculum-based measurement (CBM) data may be a helpful tool for teachers to determine the appropriate level of instruction that a child should receive. CBM can also provide information as to areas in which a child might need additional instruction and academic intervention. Instructional modifications and interventions for children may include direct instruction in particular skills that a child is lacking as well as drill and practice activities such as flashcard drills and structured worksheets.

Students with attention or behavior difficulties may also benefit from being provided with choices in academic tasks within a particular academic area. For example, students may be provided with a menu of acceptable tasks for a particular subject from which to choose. For this strategy, the teacher offers a menu of tasks that he or she finds acceptable, and the student chooses one to complete during the time allotted. The menu and choice component allows for both the teacher and the student to maintain control over the work completed. Thus, providing choices in academic tasks may increase the student's work productivity and on-task behavior.

3.2. Peer-Mediated Interventions

Peers working together on an instructional activity can also have a positive impact on the academic engagement and performance of students with AD/HD. In addition, peer tutoring strategies, by targeting academic instruction through peer interaction, are able to address both academic and social skills simultaneously. Because children with AD/HD have often been found to have difficulty in both areas, peer-mediated strategies can be a particularly useful tool.

Although different peer-mediated strategies vary regarding their foci of instruction and the structure and procedures of the tutoring teams, Barkley's 1998 handbook for diagnosis and treatment of AD/HD listed the following four common characteristics that have been shown to enhance task-related behaviors of children with AD/HD: (a) working one-on-one with another individual, (b) letting the learner determine the instructional pace, (c) continually prompting academic responses; and (d) providing frequent immediate feedback about quality of performance.

In 2002, Greenwood and colleagues described peer tutoring as an instructional strategy in which classroom teachers train and supervise students to teach their peers. Peer-directed instruction allows for high levels of student engagement, sufficient practice, immediate error correction and feedback, and the integration of students with varying abilities. In their review of peer tutoring strategies, Greenwood and colleagues described four programs that have been found to have well-defined procedures and supporting evidence in the research for use with students with disabilities: Classwide Peer Tutoring (CWPT), Peer-Assisted Learning Strategies (PALS), Classwide Student Tutoring Teams (CSTT), and Reciprocal Peer Tutoring (RPT).

3.3. Computer-Assisted Interventions

Computer-assisted interventions (CAI) can be an effective tool for promoting skill acquisition through the use of instructional technology as well as for providing additional drills and practice to promote mastery of previously acquired skills. Features of computer programs (e.g., provision of immediate performance

feedback) may be able to positively affect on-task behaviors and work production of children with AD/HD. In 2003, DuPaul and Stoner listed the following important design features of CAI that may be beneficial for children with AD/HD: specific instructional objectives presented along with activities, use of print or color to highlight important material, use of multiple sensory modalities, content divided into manageable sized chunks of information, the ability to limit nonessential distracting features (e.g., sound effects, animation), and immediate feedback on response accuracy.

3.4. Parent-Mediated Interventions

Parents should also be considered a valuable resource in promoting the academic success of their children. Frequent and clear communication between teachers and parents provides parents with knowledge about their children's academic and classroom-related behaviors and also provides teachers with information regarding events at home that may affect classroom behavior and performance. Parents can then support and reinforce instruction in the classroom as well as help to promote appropriate classroom behavior.

Parents, in providing instructional support to their children, can engage their children in flashcard drills for additional practice in an area, act as tutors to aid in skill acquisition, and make themselves available for homework support. Many parents may already participate in similar activities or would like to participate but are unsure as to how they can be most helpful. A key component to effective instructional support at home is parents' knowledge regarding instruction that their children are receiving at school, areas in which their children are struggling and could use additional support, and areas in which their children are succeeding.

Communication can take place through multiple methods. Brief notes can be sent home on a daily or weekly basis, weekly phone calls or e-mails can be scheduled, brief conversations can occur when picking up children, or another system that is mutually agreeable between families and the school can be arranged. One example of school–home communication that supports positive behavior in the classroom is a school–home contingent reinforcement system (i.e., daily report card). This type of system involves setting goals and contingencies for students. When appropriate, parents, teachers, and students should collaboratively create a template of student behaviors or goals. Approximately three to five target areas or goals should be chosen and can be rated by all teachers who are participating. Goals are individualized for each

student, should generally consist of areas of primary concern for the student, and should be stated in a positive manner. Some examples of classroom goals may include participation in class discussions, appropriate interactions with peers, work completion, and accuracy. Goals can be rated on a yes/no scale (e.g., appropriately participates in class discussion two times, completes work with 85% accuracy) or on a range of scores (e.g., a 5-point scale ranging from "excellent" to "terrible"). The report card ratings should be filled in by each teacher (in ink to prevent alterations by students) and carried from class to class by students. At home, scores on the report card can then be used in a token economy to be exchanged for privileges or points to be applied to later privileges (e.g., television time, computer time, later bedtime). Table II lists components of effective school–home communication programs.

As with classroom-based response cost systems, students should be aware of specific criteria necessary to earn privileges, which privileges they can work toward, and when they may exchange tokens for them. Exchanges for privileges might need to occur on a more frequent basis during the initial stages of a school–home contingency system and should be adapted based on the age and developmental level of each student. Motivation of the students is vital to the success of the school–home contingency system. Therefore, it may be beneficial for students to have input into goals and privileges. In addition, a space for teacher comments and parent

TABLE II
Components of Effective School–Home Communication Programs

1. Daily and/or weekly goals are stated in a positive manner.
2. Both academic and behavioral goals are included.
3. A small number of goals are targeted at a time.
4. The teacher provides quantitative feedback about student performance.
5. Feedback is provided by subject or class periods.
6. Communication is made on a regular basis (either daily or weekly).
7. Home-based contingencies are tied to school performance. Both short- and long-term consequences are employed.
8. Parental cooperation and involvement are solicited prior to implementation.
9. Student input into goals and contingencies is solicited, particularly with older children and adolescents.
10. Goals and procedures are modified as necessary.

Source. From DuPaul and Stoner (2003). Copyright 2003 by The Guilford Press. Reprinted with permission.

signatures can be provided to allow for frequent brief communications between parents and teachers. Figure 1 shows an example of what a school–home daily report card might look like.

3.5. Self-Mediated Interventions

Students can also be taught to monitor their own behavior or completion of academic-related steps. Although each student acts as the primary mediator for such strategies, it is important to note that a teacher must be present to signal recording and/or to monitor the accuracy of the student's self-recording. Checklists for work completion, or for procedures for completing tasks in specific academic areas, can be created. The presence of these checklists while the student is completing work can provide a prompt as to what steps are necessary to complete the task as well as provide for communication between the teacher and the student as to what steps have been completed.

Students can also be taught to observe and record occurrences of their own behavior—both positive and negative. A checklist or grid can be provided for each student to keep on his or her desk. The student could then record occurrences of a behavior or be taught to respond to a signal (e.g., a beep from a tape recorder) and record whether or not he or she is currently performing a specific behavior (i.e., on task) at that time. In this way, not only is the student monitoring the occurrence of his or her own behavior, but the signal also acts as a prompt to demonstrate the behavior. To be most effective, it might be advantageous to have the teacher also record occurrences. These records can then be "matched" and reinforcement can be provided based on how well the results from the teacher and the student match up. Frequency of teacher–student match checks and reinforcement schedules can then be decreased as the child becomes more adept at the procedures.

3.6. Intervention Considerations for Secondary School-Aged Students

Typically, AD/HD symptoms are chronic and can be associated with myriad additional difficulties for secondary school-aged students. Adolescents with AD/HD often exhibit problems with study and organizational skills, test performance, and social and emotional adjustment. Unfortunately, minimal research has been conducted examining school-based interventions for secondary-level students with this disorder. Although

Child's Name: _____ Date: _____					
	Special	Language Arts	Math	Reading	SS/Science
Follows class rules with no more than 3 rule violations per period	Y N	Y N	Y N	Y N	Y N
Completes assignments within the designated time	Y N	Y N	Y N	Y N	Y N
Completes assignments at 80% accuracy	Y N	Y N	Y N	Y N	Y N
Complies with teacher requests (no more than 3 instances of noncompliance per period)	Y N	Y N	Y N	Y N	Y N
No more than 3 instances of teasing per period	Y N	Y N	Y N	Y N	Y N

OTHER
Follows lunch rules   Y   N
Follows recess rules   Y   N

Total number of "yes" responses _____
Total number of "no" responses _____
Percentage _____

Teacher's initials: _____
Comments:

FIGURE 1   Sample school–home daily report card. Copyright 2001 by William E. Pelham, Jr., Ph.D. Reprinted with permission.

many of the interventions discussed for children with AD/HD may be helpful for adolescents, modifications must be made to account for expected differences in developmental level and degree of independence. For example, token reinforcement and/or response cost systems must be modified in several ways. First, the adolescent should be involved in "negotiating" the specific responsibilities and privileges that are included in the contingency management system. Second, the system will be less concrete than what typically is used with younger children. In particular, a written contract specifying reinforcement and punishment is used rather than token reinforcers such as stickers and poker chips. Third, the time period between completion of a responsibility and receipt of reinforcement might be longer than with younger children because adolescents are presumably able to handle longer delays in reinforcement. As is the case for younger students, interventions for adolescents with AD/HD should include a variety of mediators (e.g., teachers, parents, computer, peers, self), especially because secondary school students have multiple teachers who might not have the time to implement complex treatment protocols. Self-mediated interventions have particular promise for those adolescents who have demonstrated progress with other-mediated strategies and who appear to have the requisite skills for self-monitoring and self-evaluation.

In addition to the various interventions enumerated previously, secondary school students with AD/HD may require other interventions to address study and organizational skills deficits as well as difficulty in interacting with peers and authority figures. First, students can be provided with direct instruction in study and organizational skills, including taking notes, preparing for and taking tests, and completing long-term assignments. A second promising avenue for intervention for adolescents with AD/HD is the implementation of "coaching" to support students in achieving self-selected academic, behavioral, and social goals. Peers, older students, siblings, or adults can serve as coaches for students, who are assisted in setting goals, developing plans to reach stated goals, identifying and overcoming obstacles to goal attainment, and evaluating progress. Third, adolescents with AD/HD and their families are likely to require some form of counseling support to address possible difficulties in adherence with household rules as well as in interactions with family members. Support may include promoting an accurate understanding of AD/HD and its influence on families' interaction patterns; planning the education

of adolescents with AD/HD; and/or negotiating privileges and responsibilities.

4. CONCLUSIONS AND FUTURE CONSIDERATIONS

Children and adolescents with AD/HD frequently experience difficulties in school settings, most notably in the areas of academic achievement and interpersonal relationships. Thus, a comprehensive treatment plan often includes a variety of school-based interventions in combination with psychostimulant medication and home-based behavioral strategies. Effective classroom interventions include both proactive strategies that are designed to modify antecedent events and reactive strategies that focus on changing consequent events. Although treatment strategies are usually implemented by teachers, plans should include peer-mediated, self-mediated, computer-mediated, and/or parent-mediated components. Interventions that directly address the function of disruptive behavior and that are applied as close to the point of performance as possible are more likely to be effective. Students with AD/HD at the secondary school level will require additional support to enhance study and organizational skills as well as coaching to improve social relationships. Empirical investigations conducted to date have emphasized the use of reactive, consequent-based interventions. Research efforts must be directed toward establishing efficacious, antecedent-based strategies as well as toward providing more support for interventions to address the unique needs of secondary school-aged students with this disorder.

See Also the Following Articles

Attention Deficit / Hyperactivity Disorders (ADHD) ■ Child Development and Culture

Further Reading

Barkley, R. A. (1998). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford.

Dawson, P., & Guare, R. (1998). *Coaching the ADHD student*. New York: Multi-Health Systems.

DuPaul, G. J., & Eckert, T. L. (1997). School-based interventions for children with attention-deficit/hyperactivity disorder: A meta-analysis. *School Psychology Review*, 26, 5–27.

DuPaul, G. J., & Power, T. J. (2000). Educational interventions for students with attention-deficit disorders. In T. E. Brown (Ed.), *Attention-deficit disorders and comorbidities in children, adolescents, and adults* (pp. 607–635). Washington, DC: American Psychiatric Press.

DuPaul, G. J., & Stoner, G. (2003). *ADHD in the schools: Assessment and intervention strategies* (2nd ed.). New York: Guilford.

Goldstein, S., & Goldstein, M. (1998). *Managing attention deficit hyperactivity disorder in children: A guide for practitioners* (2nd ed.). New York: John Wiley.

Greenwood, C. R., Maheady, L., & Delquadri, J. (2002). Classwide peer tutoring programs. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventative and remedial approaches.* Bethesda, MD: National Association of School Psychologists.

Pelham, W. E. (2001). *Attention deficit hyperactivity disorder: Diagnosis, nature, etiology, and treatment.* (Available: Center for Children and Families, 318 Diefendorf Hall, 3435 Main Street, Buffalo, NY 14214)

Pelham, W. E., Wheeler, T., & Chronis, A. (1998). Empirically-supported psycho–social treatments for ADHD. *Journal of Clinical Child Psychology*, 27, 190–205.

Rathvon, N. (1999). *Effective school interventions: Strategies for enhancing academic achievement and social competence.* New York: Guilford.

Robin, A. L. (1998). *ADHD in adolescents: Diagnosis and treatment.* New York: Guilford.

Shapiro, E. S., & Cole, C. L. (1994). *Behavior change in the classroom: Self-management interventions.* New York: Guilford.

Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement.* New York: Guilford.

Watson, T. S., & Steege, M. W. (2003). *Conducting school-based functional behavioral assessments: A practitioner's guide.* New York: Guilford.

Attention Deficit/Hyperactivity Disorders (ADHD)

Anne Teeter Ellison

University of Wisconsin, Milwaukee, Wisconsin, USA

GLOSSARY

behavioral disinhibition Poor self-regulation or the inability to control one's activity level, attention, and emotions.

inattention Difficulties with distractibility, alertness, and arousal as well as with selective, sustained, and persistent attention to tasks.

negative parental interactions Parental interactions that are undesirable (e.g., fewer interactions, less affection, higher parental stress) or produce undesirable effects in their children.

preventive parental measures Behaviors of social control, including reinforcing prosocial behaviors, training preschool teachers, and using behavioral management principles.

psychosocial stressors Conditions in a person's social environment that affect levels of stress (e.g., poverty, family dysfunction).

self-management training Education in methods of regulating a person's behavior (e.g., self-instruction, anger control).

Attention deficit/hyperactivity disorder (ADHD) is a psychiatric disorder affecting 3 to 5% of children. Recent data suggest that the incidence of ADHD may be even higher (5–7%). ADHD is a chronic disorder, with approximately 75 to 80% of affected individuals showing evidence of significant impairment during adolescence and adulthood. Estimates suggest that 1.5 to 2.0% of adults and 2 to 6% of adolescents have ADHD. Research indicates that children with persistent ADHD have more severe symptoms, significant impairment in functioning, stressful family environments, and more adverse risk factors. There are a number of factors that interact with ADHD and further compromise adjustment over the life span.

1. CORE CHARACTERISTICS OF ATTENTION DEFICIT/ HYPERACTIVITY DISORDER

Core symptoms of inattention, impulsivity, and hyperactivity comprise the major characteristics of attention deficit/hyperactivity disorder (ADHD), with new conceptualizations emphasizing poor self-control and

behavioral disinhibition. Behavioral disinhibition is synonymous with poor self-regulation or the inability to control one's activity level, attention, and emotions. In 1997, Barkley posited that disinhibition interferes with executive control functions, including working memory, internalization of speech to guide one's behavior, motor control for goal-directed behavior, the ability to analyze and synthesize responses, and self-regulation of emotions, motivation, and arousal.

The fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) outlined three subtypes of ADHD: attention deficit disorder/predominantly inattentive type (ADD/PI), attention deficit disorder/predominantly hyperactive-impulsive type (ADD/PHI), and attention deficit disorder/combined type (ADD/C). Inattention is multifaceted and refers to difficulties with distractibility, alertness, and arousal as well as with selective, sustained, and persistent attention to tasks. Individuals with ADD/PI appear to have cognitive disabilities (e.g., spacey, "daydreamy," sluggish, easily confused) that arise from slow information processing and poor focused/selective attention. There is some debate in the literature as to whether ADD/PI is a separate disorder distinct from ADD/PHI or ADD/C. Further research is warranted on this issue, particularly as it relates to gender differences. For example, some studies report that girls have higher rates of inattention than of impulsivity and hyperactivity.

Impulsivity encompasses the inability to inhibit behavior, delay responding, and delay gratification to reach long-term goals and perform tasks. Hyperactivity is one of the more obvious core features of the disorder and refers to excessive activity (both verbal and physical). Individuals with ADD/PHI display primary problems with hyperactivity and impulsivity, whereas those with ADD/C show deficits in all core symptoms, including impulsivity, inattention, and hyperactivity.

2. ASSOCIATED PROBLEMS

The core symptoms of ADHD lead to impairment in all aspects of life activities, including school attainment, family adjustment, social relations, occupational functioning, and self-sufficiency. Sleep disorders, health problems (e.g., allergies), and accidental injuries (e.g., bone fractures, poisoning) also appear to be higher in children with ADHD than in controls. More accidental injuries may occur in youth with high levels of aggression than in those with hyperactivity alone. Long-term complications associated with ADHD include risk of substance abuse problems; problems in marriage, family cohesiveness, and chronic family conflict; employment difficulties such as frequent job changes, stress on the job, and underemployment; increased health risks such as early cigarette smoking, early sexual activity, increased driving accidents and eating disorders, and sleep disorders in children and adolescents; and risk of comorbid psychiatric disorders.

3. ADHD WITH OTHER DISORDERS

ADHD places individuals at risk for other psychiatric disorders. Oppositional defiant disorder (ODD, 54–67%) and conduct disorders (20–56%) are among the most frequent comorbid disorders in children due to impulsivity and an inability to follow rules. Conduct disorders are also common in teens (44–50%). The presence of ADHD with conduct problems increases the risk of later difficulties, including drug use and abuse, driving accidents, and additional psychiatric problems. Social problems are common in individuals with ADHD, where intrusive, inappropriate, awkward, and/or ineffective behaviors lead to rejection or strained relations with others. Mood disorders are also high (20–36%), with 27 to 30% having anxiety and 15 to 70% having major depression or dysthymia. Stimulant medication might not be as effective in individuals with ADHD plus anxiety, and side effects might be higher than in nonanxious ADHD groups. Learning disabilities are also frequent, with as many as 20 to 50% of children exhibiting significant learning problems due to difficulties with attention, work completion, and disruptive problems in the classroom.

In 1997, Jensen and colleagues cautioned that clinic-based longitudinal studies may increase the appearance of comorbidity because persons with more severe and comorbid conditions may be more likely to participate in ongoing studies. Rates of psychiatric comorbidity differ somewhat in community-based studies of children with ADHD. For example, the Multimodal Treatment Assessment (MTA) study of youth with ADHD reported comorbidity rates as follows: In the sample of children, 31.8% had ADHD alone, 33.5% had anxiety, 14.3% had conduct disorders, 39.9% had ODD, 3.8% had an affective disorder, and 10.9% had tic disorders.

Adults with ADHD have comorbidity rates of 16 to 31% for major depression, 24 to 35% for ODD, 17 to 25% for conduct disorders, 4 to 14% for obsessive–compulsive disorders, 35% for alcohol dependence or abuse, 24% for substance abuse (i.e., cannabis or other

drugs), 43% for generalized anxiety disorder, and 52% for overanxious disorder. Rates of conduct disorders or antisocial personality disorder are also high in adults with ADHD (22%).

The mechanisms of comorbidity are not well understood, but it appears that ADHD places an individual at risk for other psychiatric disorders and the presence of comorbid disorders interacts with and alters the developmental trajectory and treatment responsivity of ADHD. In 2003, Pliska and colleagues provided an assessment and treatment approach for children with ADHD and comorbid disorders. It is recommended for addressing more complex cases of ADHD.

4. GENDER AND CULTURAL ISSUES

National reports have highlighted the disparity of diagnosis and treatment for mental illness for girls and non-White children. Despite similarities in the incidence rates of hyperactivity in Black children, service delivery is lower in Black children than in Whites with similar problems. Referral and treatment rates are lower for girls (three times less likely to receive treatment compared with boys) and for Black children (three times less likely to be referred compared with White children).

There is some evidence that gender differences may be a function of the referral source, that is, clinic-referred children versus non-clinic-referred children. In a review of 18 studies on girls with ADHD in 1997, Gaub and Carlson found that nonreferred samples of girls with ADHD, as compared with boys with ADHD, had more intellectual impairments but were less aggressive, were less inattentive, and had lower levels of hyperactivity. However, when compared with clinic-referred samples, girls and boys with ADHD had more similarities than differences on core ADHD symptoms.

Other studies of clinic-referred girls indicate that girls with ADD/C are indistinguishable from boys with ADHD on measures of comorbid disorder, behavioral ratings of core symptoms, psychological functioning, and family history of psychopathology. However, when differences did occur, girls had lower reading scores and higher parent-rated measures of inattention. Furthermore, referred girls were a more extreme sample than were boys, with the former having higher rates of familial ADHD. Girls also showed a more positive response to stimulant medication than did boys. In 1999, Biederman and colleagues found similar results in girls referred to pediatricians and psychiatrists. Girls referred for ADHD were more likely to show conduct problems, mood and anxiety disorders, a lower intelligence quotient (IQ), and more impairment on social, family, and school functioning than were nonreferred girls. Conduct problems were lower in girls than in boys with ADHD, and this may account for lower referral rates. However, high rates of mood and anxiety disorders suggest the need for comprehensive treatment. It was of particular note that girls in this study also had high rates of substance abuse disorders, including alcohol, drug, and cigarette use, and were at an increased risk for panic and obsessive–compulsive disorders.

In sum, clinic-referred girls with ADHD present with more symptoms than do girls without ADHD and are indistinguishable from clinic-referred boys with ADHD. It is important to note that girls who are referred for ADHD may represent a more severely impaired group than do community-based samples of girls. Further evidence suggests that girls who are in need of treatment for ADHD may be overlooked because they tend to be less disruptive than boys. Additional longitudinal research investigating gender differences would be helpful to resolve these critical issues.

5. ETIOLOGY OF ADHD

Research indicates that problems in behavioral inhibition or self-control are a result of dysfunction in frontal–striatal networks, whereas other brain regions (e.g., basal ganglia that includes the caudate nucleus and cerebellum) are also implicated. The evidence of genetic transmission of ADHD, primarily involving the dopamine systems that innervate frontal–striatal regions, is strong. Studies estimate that 70 to 95% of deficits in behavioral inhibition and inattention are transmitted genetically. Research investigating the manner in which the environment interacts with subtle brain anomalies and genetic mutations is ongoing. Traumatic events, the presence of comorbid disorders, and other psychosocial stressors (i.e., poverty, family dysfunction) complicate ADHD but are not considered to be causal.

Although neurological and genetic substrates appear to be compromised, multiple interacting factors are likely involved in the expression of ADHD. It is likely that compromised neural systems influence adaptive functioning and that family, school/work, and community environments affect how ADHD is manifested and may contribute to the development of various coexisting disorders. There may be other factors, including exposure to environmental toxins (e.g., elevated lead exposure), prenatal smoking, and alcohol use, that

increase the risk for ADHD. However, these risk factors are not present in all children with ADHD. Some environmental explanations are inaccurate and non-scientifically based, including high sugar ingestion, allergies or sensitivities to foods, family discord, parental alcoholism, poor or ineffective parenting, and poor motivation. Furthermore, inaccurate beliefs about the nature of ADHD often lead to ineffective treatment approaches.

6. DEVELOPMENTAL CONTEXT FOR ADHD

Although ADHD has been considered to be a disorder of childhood, there is compelling evidence that a majority of children do not outgrow ADHD. This section highlights the major challenges of ADHD throughout the life span.

6.1. ADHD During Early Childhood

Symptoms of ADHD typically first appear during early childhood. Infants are often described as temperamental, difficult to care for due to excessive crying and irritability, difficultly in calming, overly sensitive to stimulation, and overly active. These complications often interfere with normal parent–child bonding and often lead to negative parental interactions (e.g., fewer interactions, less affection, higher parental stress). These patterns affect child compliance and lead to frustrating and challenging interactions. In preschool, hyperactivity levels are pronounced and lead to difficulty in adjusting to expectations to sit, listen, and get along with other children. Impulsivity interferes with play and often leads to rejection. Referral rates are high during this stage of development as children come into contact with other adults and face greater demands for self-control. Interventions at the stage frequently focus on increasing parenting skills, building positive parent–child relationships, increasing parent support, and implementing other preventive measures (e.g., reinforcing prosocial behaviors, training preschool teachers, using behavioral management principles).

6.2. ADHD During Middle Childhood

Most research on ADHD has been conducted on children between 6 and 12 years of age. Deficits in self-control continue to be problematic and are highlighted by disruptive, noncompliant, and off-task behaviors at home and at school. Poor attention to schoolwork, poor work completion, low motivation, low persistence to challenging tasks, and poor organizational skills negatively affect academic and school adjustment. Difficulties with social situations, negative peer and adult interactions, poor anger control, and low self-esteem create secondary problems that can be chronic. Parent–child relationships are often strained due to noncompliance, failure to complete household chores, and the need for constant monitoring of everyday activities (e.g., bathing, eating, getting dressed, going to bed). These difficulties can increase family stress and interfere with sibling relationships. Other comorbid disorders, including oppositional deviance, conduct problems, depression, and anxiety, may also emerge during this stage. Severe oppositional deviance is problematic and often presages antisocial behavior during later adolescence and adulthood. Treatments during this stage typically are multimodal, including parent training, behavior classroom management, academic interventions, self-management training (e.g., self-instruction, anger control), and medication.

6.3. ADHD During Adolescence

Longitudinal studies reveal significant difficulties for approximately 70 to 80% of teens who had ADHD as young children. Although cognitive deficits and learning disabilities are common in children with ADHD, they are less well documented in adolescents and adults. Longitudinal studies show that youth with ADHD have significant academic difficulties, including high suspension rates (46%), high dropout rates (10%), and placements in special education for learning disabilities (32.5%), emotional disturbance (35.8%), and speech language disorders (16.3%). Negative academic outcomes were present even after intensive treatments, including medication, individual therapy, family therapy, and special education placement. In general, children with hyperactivity are less well educated, have higher rates of grade retention, and have lower grades compared with controls at 5- and 10-year follow-ups. Adolescents with ADHD also show higher rates of automobile accidents and speeding tickets than do teens without ADHD. Both cigarette use and marijuana use are higher according to parental reports, whereas teens with ADHD report higher rates of cigarette use but not of alcohol use, or of marijuana, cocaine, heroin, and other illegal substances use, compared with non-ADHD teens. Antisocial behaviors, including theft, breaking and entering, disorderly conduct, carrying a weapon, assault with a weapon, assault with fists, setting fires,

and running away from home, were reported in the Milwaukee Longitudinal Study.

In general, treatment options for teens are less well researched than are those for children. Treatment for adolescents with ADHD typically focuses on increasing problem solving and communication between parents and teens, psychopharmacotherapy, and classroom accommodations for academic difficulties. More systematic study is needed to investigate the strength of these various interventions. It has been suggested that children with ADHD might not receive needed treatment.

6.4. ADHD During Adulthood

In 1996, Barkley and colleagues found that adults with ADHD had similar levels of educational achievement and occupational adjustment but differed from controls on symptoms of ADHD and oppositional problems in college and at work. They also had shorter duration of employment, more psychological distress and maladjustment, and more antisocial acts and arrests for disorderly conduct and thefts compared with controls. Even though conduct problems and risk for comorbid antisocial personality disorder appear in approximately 25% of individuals with ADHD, the majority of adults with hyperactivity do not engage in criminal behaviors.

According to driving instructors and self- and parent-reports of driving skills, young adults with ADHD were more distractible and impulsive while driving. High rates of driving-related difficulties, including license suspensions or revocations, serious accidents (i.e., involving a wrecked car), and hit-and-run accidents, were also reported. In 1998, Barkley indicated that young adults with ADHD had sexual intercourse at an earlier age, more sexual partners, and higher rates of pregnancy. Contraceptive use was lower, sexually transmitted diseases were higher, and testing for HIV/AIDS was higher in the ADHD group than in controls.

Treatments for adults with ADHD are not well documented but often include multiple approaches, including family and couples therapy, occupational and career counseling, occupational accommodations, medication, and treatment for comorbid disorders (e.g., alcohol or drug treatment, depression, bipolar disorders). Others have emphasized the need for counseling to change the negative mind-set that results from years of failure and coaching for everyday responsibilities. Although empirical studies are needed to determine the efficacy of various treatment options, studies do show the efficacy of stimulant medication in the treatment of ADHD in adults.

7. EVIDENCE-BASED INTERVENTIONS FOR ADHD

The American Academy of Child and Adolescent Psychiatry (AACAP) has developed practice guidelines for the diagnosis and treatment of ADHD. The American Academy of Pediatrics (AAP) recommends that stimulant medication should not be used as the only treatment for ADHD and should be administered only after a careful evaluation. Practice guidelines recommend a comprehensive multimethod approach for the diagnosis of ADHD in children and youth.

7.1. Multimodal Treatment Regimens

The MTA study, funded by the National Institute for Mental Health, reported that children with ADHD received suboptimal care in the community. Even though two-thirds of the sample received stimulant medication, care in the community was less effective than were carefully managed medication, behavioral treatment, and combined treatments. Only 25% of children receiving care in the community were normalized after a 14-month trial, whereas 68% of the combined group, 56% of the carefully managed medication group, and 34% of the behavioral treatment group showed normalization. When community care is provided, it is not carefully monitored, nor is it as effective as multimodal intensive treatment. The MTA study showed that children with ADHD had the best response to multimodal treatment that included 35 sessions of parent training, a full-time summer treatment program for children to learn social and sports skills and to practice academic skills, weekly teacher consultation, a paraprofessional aide in the classroom, contingency management in the classroom, and medication. This extensive treatment was highly effective in reducing the major symptoms of ADHD and was superior to treatment generally found in the community. It is difficult to discern whether the quality and level of treatment described in the MTA study can be easily implemented in the community.

7.2. Other Evidence-Based Interventions

Other empirically supported treatments for children with ADHD include behavioral therapy and contingency management techniques; a summer treatment program with a systematic reward/response cost program, sports skills

training, a 1-hour daily academic special education class, training in effective social skills, daily report cards, and parent training; parent training combined with contingency management and didactic counseling to increase parent knowledge of ADHD; a community-based family therapy program; the good behavior game, response cost, using the "attention trainer"; modification of classroom assignments and task demands; and the Irvine Paraprofessional Program. Self-management, direct contingency management, and intensive behavioral and social skills training have also been shown to be effective.

8. MEDICATION MONITORING AND ADHERENCE

Research on the short-term efficacy of stimulant medication is well documented for 75 to 80% of children with ADHD; however, medication monitoring and adherence is problematic. For example, in 2001, Vitiello stated, "For optimal pharmacological treatment of children with ADHD, medication adjustments are needed for long-term treatment even when the initial dose is chosen in a careful, comprehensive, and unbiased manner." In the MTA study, more than 70% of children assigned to the medical management group were on different doses after a 13-month trial. Although the majority of children receiving community care in the MTA study were treated with stimulant medication, as a whole, children in the medical management group received more careful medication monitoring from physicians and showed greater improvement of symptoms than did children in the care in the community group.

In 2001, Thiruchelvam and colleagues investigated medication adherence in children 6 to 12 years of age. The most salient factors affecting compliance were the absence of ODD, the severity of ADHD symptoms, and the age of the children. Children with ODD were 11 times more likely to refuse medication. Youth with more symptoms were more responsive to the medication and were more compliant when taking medication as well. Positive stimulant response may encourage parents and children to stick to the medication regimen at higher rates than in cases where medication is not very helpful. Older children were also more likely to refuse medication. Because there is a decrease in hyperactivity symptoms with age, older children may perceive less benefit from medication and choose not to adhere. Social stigma may also play a role in adherence at this stage. Physicians are advised to develop adherence plans for youth on medication.

In sum, research indicates that multiple therapies are needed to adequately address the problems associated with ADHD. Stimulant medication with behavioral and psychosocial interventions, including classroom behavior management and parent management training, improve ADHD symptoms and associated problems in children with ADHD. Initial research indicates that stimulant medication is effective for adults and that other cognitive–behavioral interventions (e.g., self-management) show promise. Currently, treatment within a single modality (medication vs behavioral) appears to have very little long-term impact. Effective treatments are less well documented for adolescents and adults with ADHD. Although studies have shown that stimulant medication is effective, less is known about the effects of multimodal treatment in older groups.

9. RISK AND RESILIENCY FACTORS THAT AFFECT ADHD

Although ADHD presents challenges throughout the life span, some factors complicate the disorder, whereas others appear to be protective. Risk factors that alter the course of ADHD include child characteristics (e.g., severity of ADHD symptoms, intelligence levels, comorbidity), family discord or environmental distress, and early treatment for ADHD and the presence of coexisting disorders. In an effort to optimize outcome, comorbid disorders should be targeted for treatment along with the ADHD symptoms. Oppositional defiant behavior problems are among the most debilitating difficulties over time because they often lead to conduct disorders and antisocial personality disorders during adolescence and adulthood. There is evidence that effective parenting skills can interrupt this progression in many children. Other family factors that increase the complexity of ADHD include parental psychopathology such as maternal depression and paternal antisocial personality disorder. These parental difficulties often interfere with effective parenting and produce added stress to vulnerable families whose members are already challenged by disruptive noncompliant child behaviors. There is strong evidence that raising a child with ADHD is stressful, so additional parental problems make this challenge even more overwhelming. Furthermore, as noted by Goldstein in 2002, "Living in a household, above the poverty level, with parents who are free of serious psychiatric problems, consistent in their parenting style, and available to their children appear to be among the most

powerful variables at predicting good outcome." Parents are also advised to seek individual and family therapy for their own problems in an effort to strengthen interpersonal effectiveness and family cohesiveness.

The extent to which treatment alters the developmental course of ADHD is not well understood. Some studies of children who received extensive treatment showed that these youth still had poor outcomes during adolescence. However, the MTA study showed short-term improvement (over a 14-month period) in children receiving comprehensive multimodal treatment in highly controlled and monitored programs. Furthermore, growing evidence suggests that medication may buffer some of the negative effects of ADHD. Although there are reasons to be cautious about the use of stimulant medication in young children, recent studies suggest that early treatment may alter the neurodevelopmental pathways of ADHD in positive ways. Stimulant treatment also appears to improve outcomes for adults. In 1999, Wilens and colleagues also found that 70% of adults receiving a year of cognitive therapy and stimulant medication showed a reduction of ADHD symptoms and were less anxious and depressed.

Current studies of ADHD are focusing on the long-term effects of multimodal treatment for children and are investigating what works best for adolescents and adults with ADHD. Current research is promising, but researchers are still exploring how the outcome and course of ADHD can be altered with effective treatments. The impact of environmental events on the development of attention and self-regulation is of particular interest.

See Also the Following Articles

Child Development and Culture ■ Educational and Child Assessment ■ Family and Culture

Further Reading

Barkley, R. A. (1997). *ADHD and the nature of self-control.* New York: Guilford.

Barkley, R. A. (1998). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* New York: Guilford.

Barkley, R. A., & Gordon, M. (2002). Research on comorbidity, adaptive functioning, and cognitive impairments in adults with ADHD: Implications for a clinical practice. In S. Goldstein, & A. Teeter Ellison (Eds.), *A clinician's guide to adult ADHD: From evaluation to treatment* (pp. 43–69). San Diego: Academic Press.

Barkley, R. A., Murphy, K., DuPaul, G., & Bush, T. (2002). Driving in young adults with attention deficit hyperactivity disorder: Knowledge, performance, adverse outcomes, and the role of executive functioning. *Journal of the International Neuropsychological Society*, 8, 655–672.

Castellanos, F. X., Lee, P., Sharp, W., Jeffries, N., Greenstein, D., Clasen, L., Bluementhal, J., James, R., Ebens, C., Walter, J., Zijdenbos, A., Evans, A., Giedd, J. N., & Rapoport, J. L. (2002). Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Medical Association, 288*, 1740–1748.

DuPaul, G., & Stoner, G. (2003). *ADHD in the schools: Assessment and intervention strategies* (2nd ed.). New York: Guilford.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the Academy of Child and Adolescent Psychiatry*, 36, 1036–1045.

Goldstein, S. (2002). Continuity of ADHD in adulthood: Hypothesis and theory meet reality. In S. Goldstein, & A. Teeter Ellison (Eds.), *Clinician's guide to adult ADHD: Assessment and intervention* (pp. 25–45). London: Academic Press.

Leibson, C. L., Katusic, S. K., Barbaresi, W. J., Ransom, J., & O'Brien, P. C. (2001). Use and cost of medical care for children and adolescents with and without attention-deficit/hyperactivity disorder. *Journal of the American Medical Association*, 285, 60–66.

MTA Cooperative Group. (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, 56, 1073–1086.

Pliszka, S. R., Carlson, C. L., & Swanson, J. (2003). *ADHD with comorbid disorders: Clinical assessment and management.* New York: Guilford.

Tannock, R. (1998). Attention deficit hyperactivity disorder: Advances in cognitive, neurobiological, and genetic studies. *Journal of Child Psychology and Psychiatry*, 39, 65–100.

Teeter, P. A. (1998). *Interventions for ADHD: Treatment in developmental context.* New York: Guilford.

Teeter Ellison, P. A. (2002). An overview of childhood and adolescent ADHD: Understanding the complexities of development into the adult years. In S. Goldstein, & A. Teeter Ellison (Eds.), *Clinician's guide to adult ADHD: Assessment and intervention* (pp. 2–19). London: Academic Press.

Thiruchelvam, D., Charach, A., & Schachar, R. J. (2001). Moderators and mediators of long-term adherence to stimulant treatment in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 922–928.

Vitiello, B. (2001). Psychopharmacology for young children: Clinical needs and research opportunities. *Pediatrics*, 108, 983–989.

Wilens, T. E., McDermott, S. P., Biederman, J., & Abrantes, A. (1999). Cognitive therapy in the treatment of adults with ADHD. *Journal of Cognitive Psychotherapy*, 13, 215–226.

Attitude Measurement

David Trafimow

New Mexico State University, Las Cruces, New Mexico, USA

GLOSSARY

accessibility The ease with which a memory can be activated; it is often measured by the time it takes a person to react to a stimulus.

action What a person is actually going to do or not do; it is usually the verb in the attitude measure.

associative hypothesis The hypothesis that a person forms associations between behavioral beliefs and other behavioral beliefs, between normative beliefs and other normative beliefs, but not between behavioral beliefs and normative beliefs.

construct validity The extent to which a measure of a variable actually measures what it is supposed to measure.

context The situation in which a person will perform or not perform the behavior of concern.

factor analysis A statistical technique that reduces a large number of test items or scales to a smaller number of dimensions.

internal consistency The extent to which different scales that comprise a measure correlate with each other.

open measurement A way in which to measure repeated behaviors that may increase the validity of the measures of corresponding attitudes.

predictive validity The extent to which a measure of a variable predicts the measure of another variable.

principle of correspondence The extent to which an attitude and behavior measure match each other with regard to action, target, time, and context.

repeated behavior A behavior that is performed several times or more; such a behavior presents a challenge for the valid measurement of its corresponding attitude.

skewed distribution A distribution that deviates from normality as a result of more extreme scores at one tail than at the other tail; response latencies are notorious for having extreme scores at the tail that represents slow latencies (i.e., long times).

target The object that is or is not acted on by a person whose attitude is being measured.

test-retest reliability The correlation between a test and itself when it has been taken at two different times.

time When the behavior will or will not be performed by a person whose attitude is being measured.

Because attitudes have been assumed to be predispositions for behaviors, consumer researchers

have assumed that attitudes should predict behaviors. However, the success of predicting behaviors from attitudes largely depends on how reliably and validly attitudes are measured. The validity of attitude measures, in turn, depends on a variety of issues, the most important of which is correspondence of measurement. Other concerns are whether one is measuring attitudes toward objects or toward behaviors, whether one wishes to obtain implicit or explicit attitude measures, whether the attitude pertains to a single behavior or to repeated behaviors, and whether attitudes can be distinguished from other variables.

1. INTRODUCTION

The vast majority of applied researchers who measure attitudes do so because they wish to predict or affect people's behaviors. Health psychologists wish to increase healthy behaviors such as exercise, healthy eating, and use of condoms; consumer researchers wish to induce people to buy products; and politically motivated researchers wish to predict and affect people's votes and financial contributions. Researchers usually assume that attitudes are an important determinant of behaviors, with the accompanying assumption that knowledge of people's attitudes implies an ability to predict and control behaviors. To obtain such knowledge of people's attitudes, however, it is absolutely crucial to be able to measure them reliably and validly—the main topic of this article.

2. TWO MEANINGS OF RELIABILITY

An attitude measure can be reliable in one of two general ways. To understand the simplest type of reliability, suppose that a set of participants completes an attitude measure at a particular time and then completes it again at a later time. The correlation between the two test-taking occasions is the attitude measure's test–retest reliability. However, there is a second meaning of reliability in that the term can be used to denote internal consistency. To understand internal consistency, suppose that an attitude measure is composed of several scales. If these scales truly measure the participants' attitudes, all of the scales should correlate with each other. Another way of saying that the scales correlate with each other is to say that they have a high degree of internal consistency. Although there are different formulas for calculating different kinds of internal consistency, all of these formulas

require an assumption that all of the scales that compose the measure are equally good. Based on this assumption, all of the formulas imply that increasing the number of scales increases the internal consistency of the whole attitude measure. Because both test–retest reliability and internal consistency are usually considered to be prerequisites for validity, it is common practice to use several scales to measure attitudes.

3. TWO MEANINGS OF VALIDITY

How does one know whether a particular attitude measure is valid? To answer this question, one must first ask "valid for what?" An attitude measure can be valid if it predicts another variable that is considered desirable to predict such as behaviors and behavioral intentions (i.e., motivations to perform behaviors). For example, a measure of condom use attitudes that actually predicts whether people will intend to use condoms, or whether they actually use condoms, can be said to have predictive validity.

Another answer to the "valid for what?" question is that the measure must actually measure attitude and not something else. It is possible that a particular attitude measure might predict behavior, not because it measures attitude but rather because it measures something else that is correlated with behavior. Researchers who are concerned with this are forced to deal with the issue of how to define an attitude. Not surprisingly, various researchers have used different definitions, depending on their theories. From this perspective, the valid measurement of attitudes depends on the theories in which attitudes play an important role. Empirically obtained evidence in support of a particular theory also supports the validity of the attitude measurement as it is specified by that theory, whereas empirical evidence against a particular theory tends to disconfirm the validity of the attitude measurement as it is specified by that theory. When the theory in which the attitude construct is embedded is empirically supported, the attitude measure that is part of that theory is said to have construct validity.

4. THE HISTORICAL PROBLEM: ATTITUDES DID NOT PREDICT BEHAVIORS

Since the 1920s, researchers have assumed that attitudes are predispositions for behaviors. If this is true, it

follows that attitudes should predict behaviors. However, the vast majority of researchers from the 1920s through the 1960s obtained correlations that were either low or not statistically significant. The consistent inability of researchers to predict behaviors from attitudes suggested one of two possibilities. First, perhaps attitudes really do not matter very much for predicting behaviors. Second, perhaps attitudes do matter, despite the lack of empirical support, but the attitude measures used in previous studies were not valid. Although many researchers argued for the first possibility, other researchers eventually achieved a good deal of success by assuming the second possibility. These latter researchers redefined the idea of attitude in two ways, both of which resulted in attitude measures with greater construct and predictive validity.

4.1. Attitude Accessibility

One view derives from the social cognition tradition, but it does not receive much attention here because it has not been used much by applied researchers. Fazio used the idea of attitude accessibility to explain low correlations between attitudes and behaviors. He assumed that attitudes affect behaviors, but only when they are easy to retrieve from memory. When attitudes are not accessible, there is no reason to believe that they would affect behavior. Therefore, researchers who subscribed to this view explained previous findings of low attitude–behavior correlations by assuming that they were due to a failure on the part of researchers to measure accessible attitudes. In support of the accessibility view, much research during the 1980s and 1990s indicated that accessible attitudes are more predictive of behaviors than are less accessible attitudes.

Nevertheless, applied researchers have, for the most part, ignored the accessibility view. The reason for this is not clear. One possible reason is that many applied researchers are not particularly knowledgeable about this view. Another possible reason is that researchers who have favored the accessibility view have not been clear about how this view can be applied to predict and control behaviors.

4.2. Principle of Correspondence

Fishbein proposed the principle of correspondence, which provided the other solution to the problem of low attitude–behavior correlations. His main assumption is that behaviors have four components: action, target, time, and context. It is easiest to understand this idea with an example. Suppose that a researcher wishes to predict whether people will give blood at the campus blood drive on Tuesday. The action is "give," the target is "blood," the time is "on Tuesday," and the context is "at the campus blood drive." If the attitude measure does not correspond with the behavior measure in regard to action, target, time, and context, it is invalid and a respectable attitude–behavior correlation is not likely to be obtained. For example, if a researcher measures people's attitudes toward "giving blood," there is no reason to expect that this attitude measure will successfully predict whether people will "give blood at the campus blood drive on Tuesday." Rather, to predict this behavior, it is necessary to measure people's attitudes toward "giving blood at the campus blood drive on Tuesday." This is often done by having people make a check mark on a scale with various options such as the following:

> I extremely like/quite like/slightly like
> /neutral/slightly dislike/quite dislike
> /extremely dislike giving blood at the
> campus blood drive on Tuesday.

(Whether people actually performed the behavior could be assessed in a variety of ways such as checking the lists of people who donated at the campus blood drive on Tuesday and simply asking people whether they had given blood at the campus blood drive on Tuesday.)

The correspondence view explained previously obtained low attitude–behavior correlations in a straightforward way. The reason for the low correlations is that the attitude measures used in those studies were not valid because they did not obey the principle of correspondence. Results from two types of research paradigms supported this view. First, researchers measured attitudes and behaviors (or behavioral intentions, which are often assumed to be direct precursors to behaviors) according to the principle of correspondence and obtained much higher correlations than those that researchers had been able to obtain previously. Second, some researchers manipulated the degree of correspondence of attitude and behavior (or behavioral intention) measures experimentally and obtained much higher correlations when there was a high degree of correspondence than when there was not. In combination, the findings provided a very convincing case for the importance of measuring attitudes in accordance with the

principle of correspondence. Therefore, it is now common practice for applied researchers to measure attitudes in this way, and the result has been a dramatically improved ability to predict behavior in a variety of domains such as dieting, exercising, drinking, smoking, voting, using seat belts, using condoms, and getting screened for cervical cancer.

5. FACTOR ANALYSIS

As was discussed earlier, increasing the number of scales can increase the internal consistency of the attitude measure. Consequently, many researchers do not stop at having one attitude scale and instead have several scales. For example, in addition to the like–dislike example given previously, participants could respond to scales that include pairs such as wise–foolish, beneficial–harmful, enjoyable–not enjoyable, good–bad, and pleasant–unpleasant. When multiple scales are used to measure an attitude, the mean of a participant's responses to the scales is often taken as the representation of his or her attitude.

Depending on the behavior of interest, it may happen that not all of the scales are equally good for measuring attitude. It may even turn out that, for some behaviors, some of the scales will measure something other than attitude. To test this possibility, many researchers habitually submit the various scales to a factor analysis. Factor analysis is a statistical technique that reduces a large number of items down to a smaller number of underlying dimensions. For example, on intelligence quotient (IQ) tests, a large number of items may be reduced down to dimensions such as verbal ability and mathematical ability. As another example, on personality tests, hundreds of items have been factor analyzed and reduced to five basic factors of personality. Ideally, when various attitude scales are submitted to a factor analysis, one factor that represents attitude should result. Although one factor is often obtained, it sometimes happens that more than one factor is obtained. In this case, the researcher must determine which factor is the "true" attitude factor—a determination that depends on a variety of issues that are too complicated and numerous to discuss here. In addition, obtaining more than one factor can be interpreted to mean that there is more than one component to the attitude. In this case, the answer to the question of which factor represents attitude is that all of them may do so.

6. ATTITUDES TOWARD OBJECTS OR BEHAVIORS AND BEHAVIORAL CATEGORIES

The principle of correspondence implies a distinction between attitudes toward objects and attitudes toward behaviors. Suppose that a researcher is interested in using advertising to reduce discrimination and wishes to measure relevant attitudes. Unfortunately, it is not clear what the relevant attitudes are. In general, researchers in this area have measured attitudes toward the groups of interest such as women, African Americans, Jews, handicapped people, and people with AIDS. However, the principle of correspondence suggests that if one wishes to predict discriminatory behavior, this is not going to work. This is because although a measure of attitudes toward a particular group includes the target component of the behavior to be predicted, it does not include the action, time, and context components. If one wishes to predict a particular discriminatory behavior, it is necessary to use an attitude measure that includes all of the components of the criterion behavior. At the very least, this implies that the attitude measure will have to include the action component, which makes it a measure of an attitude toward a behavior rather than a measure of an attitude toward an object. For example, an attitude toward "hiring women" is an attitude toward a behavior, whereas an attitude toward "women" is an attitude toward an object. For a more mundane example, one could measure an attitude toward "Skippy peanut butter" or toward "buying Skippy peanut butter"; the former is an attitude toward an object, whereas the latter is an attitude toward a behavior.

An argument has been made against using the principle of correspondence in some situations. In brief, the argument is that researchers are sometimes interested in a large number of behaviors. For example, a researcher might be interested in a wide range of behaviors that discriminate against women and not just hiring behaviors. Obviously, an attitude measure toward a single behavior is likely to be inadequate for predicting a variety of discriminatory behaviors. One solution to this problem is for the researcher to measure attitudes toward all of the discriminatory behaviors that are of interest and to predict each discriminatory behavior from its corresponding attitude measure. Another solution is for the researcher to list a set of discriminatory behaviors of interest and then lump all of them under the general category of "discriminatory behaviors." Participants can then give their attitudes toward the

whole behavioral category, and these can then be used to predict the behaviors in that category. Although there is some evidence that either of these solutions may have some validity, there is insufficient evidence for a strong conclusion. However, there is strong evidence that attitudes toward behaviors are better predictors of behaviors than are attitudes toward objects.

7. ATTITUDE SPECIFICITY, BEHAVIORAL INTENTIONS, AND BEHAVIORS

Attitudes have been shown to be good predictors of behaviors, but they have also been shown to be good predictors of people's intentions to perform behaviors. In addition, such behavioral intentions have often been shown to be good predictors of behaviors, and most researchers believe that intentions are proximate causes of behaviors. This combination of empirical findings and theorizing in the area, along with the practical point that behavioral intentions tend to be much easier to measure than are real behaviors, has resulted in the widespread use of behavioral intentions as a substitute for behaviors. This substitution has resulted in two controversies that have not yet been settled. First, there is a controversy about whether behavioral intentions are close enough to real behaviors to justify using the former as a substitute for the latter. A large part of this controversy is based on the issue of how well behavioral intentions predict behaviors. To the extent that intentions do a good job of predicting behaviors, the substitution would be supported, whereas a lack of prediction would support the reverse conclusion. In fact, results vary widely. In general (but there are exceptions), behavioral intentions do a better job of predicting behaviors when the intention and behavior measures conform to the principle of correspondence than when they do not. Thus, the substitution of behavioral intentions for behaviors is most acceptable when one uses correspondent measures.

The use of behavioral intentions as a criterion measure brings up a second controversy. To understand the underlying reason for the controversy, consider that all behaviors are performed with a target, an action, a time, and a context, meaning that the measured behavior will automatically have these four elements. Therefore, the trick to obeying the principle of correspondence is to make sure that the attitude measure specifies all of these elements in a way that matches the behavior measure. But matters change when attitudes are used to predict

behavioral intentions rather than actual behaviors. A behavioral intention does not necessarily have a target, an action, a time, and a context, as is illustrated by the following example. Imagine the behavior of "buying a television set." For a person to actually perform this behavior, he or she must buy a particular brand of television set, at a particular time, at a particular store or from a particular Web site. But for this person to intend to perform the behavior, many of these elements need not be specified. In the case of this example, the intention specifies a rather vague target (television set rather than a specific brand of television set) and an action (buy) but not a time or a context. Therefore, to have an attitude measure that is correspondent with the intention measure, it too should specify the target (in a similarly vague way) and the action but not a time or a context. Consequently, when one attempts to predict behavioral intentions from attitudes, it is possible to fully specify target, action, time, and context or to not fully specify these four elements. In either case, the principle of correspondence is obeyed so long as the degree of specification of the four elements is the same for both the attitude and behavioral intention measures. Does the degree of specification affect the size of the obtained correlation between the attitude and behavioral intention measures? Currently, there is insufficient empirical evidence to answer this question. Moreover, even for the prediction of actual behaviors, it is possible for the researcher to be uninterested in a particular element (e.g., the store at which the television set was bought). In this case, the behavior measure can ignore the uninteresting element or not ignore it, and the attitude measure can ignore it or not ignore it, and so long as the two measures match, the principle of correspondence is obeyed. Will the prediction of the behavior from the attitude measure be greater if the uninteresting element is specified? Again, there is insufficient empirical evidence to know. In sum, whether one predicts behavioral intentions or behaviors, whether one predicts clear-cut behaviors or behavioral categories, and whether one is interested in more of the elements (e.g., target, action, time, context) or fewer of them, there is insufficient research to determine whether it is better to have measures that are as specific as possible or not.

8. OPEN ATTITUDE MEASURES

Although there is widespread agreement that it is crucial to measure attitudes in accordance with the principle of

correspondence, it has not always been clear how to do this. In particular, it has not always been clear how to measure attitudes toward repeated behaviors in accordance with this principle. To see the problem, suppose that a researcher is interested in predicting exercise behaviors from attitudes toward exercising. If the researcher were interested in only one performance of an exercise behavior, it would be easy to obey the principle of correspondence as follows:

Attitude measure: I like/dislike to engage in vigorous physical activity at least one time during the month of October. (Participant makes a check mark on a scale.)

Behavior measure: I engaged in physical activity at least one time during the month of October. (Participant confirms or disconfirms that he or she performed the behavior.)

Unfortunately, this method of measuring attitudes and behaviors does not work well for repeated behaviors. Suppose that a researcher is interested in regular exercise. There is no clear dividing line between "regular" and "not regular" exercise; therefore, it is not satisfactory to arbitrarily specify a number of exercise occasions that qualifies as regular. Courneya proposed the idea of open attitude and behavior measures to address this problem. His idea was to have the participant, rather than the researcher, specify the number of behaviors. Although the use of open measures has not been tested exhaustively (and the tests have focused more on predicting behaviors from intentions than from attitudes), preliminary evidence suggests that it results in an increase in the prediction of behaviors compared with other methods. The following is an example of open measures:

Attitude measure: I like to engage in vigorous physical activity _____ times during the month of October.

Behavior measure: I engaged in vigorous physical activity _____ times during the month of October.

To see why open measures often improve the prediction of behaviors from attitudes, consider an example. Suppose that a person believes that exercising 16 days during the month of October is "regular exercise," whereas exercising 15 days or less is not. In addition, suppose that the person has a positive attitude toward regular exercise but that he or she exercised only 15 days during October. With traditional measures, the person's behavior would seem to be inconsistent with his or her attitude; exercising only 15 days is inconsistent with a positive attitude toward regular exercise

(16 days). In contrast, using open measures, it can easily be seen that exercising for 15 days is quite consistent with the person's positive attitude toward exercising for 16 days.

9. IMPLICIT ATTITUDE MEASURES

Attitude researchers have often found it useful to distinguish between implicit and explicit attitude measures. Explicit attitude measures are usually characterized as being conscious, deliberative, and controllable, whereas implicit attitude measures are usually characterized as being unconscious, unintentionally activated, and not controllable. Whether the distinction between explicit and implicit measures should be considered to be a dichotomy or a continuum is not a settled issue, although many researchers who favor implicit measures seem to favor a dichotomous interpretation.

There are several examples of implicit attitude measures. One kind of implicit measure makes use of response latency, where participants respond to a computer-presented stimulus by pressing a key, and faster reaction times are interpreted to indicate a stronger attitude toward the stimulus. Other kinds of implicit measures include memory tasks, physiological measures (e.g., galvanic skin response, heart rate), and indirect questionnaires (where participants are not asked directly about the attitude object or behavior).

Implicit measures, when compared with explicit ones, have both advantages and disadvantages. Some advantages are as follows. First, if the issue of interest is a socially sensitive one, it may be difficult to get honest responses through explicit measures, and so implicit measures provide a way in which to avoid this difficulty. A finding that supports this argument is that implicit and explicit attitude measures have been found to be more highly correlated for socially insensitive attitude objects than for socially sensitive ones. Second, if the researcher is interested in unconscious attitudes, a case can be made that implicit measures are more valid than explicit ones. Supporting evidence indicates that implicit attitude measures tend to be more predictive of implicit measures of other variables than do explicit attitude measures. Third, it is possible that implicit measures are more direct than explicit ones. This is because explicit attitude measures must be funneled through the conscious processing system, whereas implicit ones do not.

On the other hand, implicit measures also have disadvantages. First, if one is interested in predicting

behavior, it is reasonably clear how to use the principle of correspondence to obtain valid explicit attitude measures, whereas it is less clear how the principle of correspondence can be used for implicit attitude measures. Second, the reliability of implicit attitude measures has not been thoroughly investigated, and the preliminary evidence suggests that they have less test–retest reliability than do explicit measures. Because reliability is a precursor to validity, this preliminary evidence also suggests that implicit measures may also be less valid than explicit ones. Third, some of the same studies that suggest that implicit attitude measures are better than explicit ones for predicting implicit measures of other variables also suggest that implicit attitude measures work less well than explicit ones for predicting explicit measures of other variables (e.g., behavioral intentions). Finally, the presumed advantages for implicit attitude measures depend on assumptions about the nature of consciousness, cognitive resources, controllability, and others that, although supported in the literature, have not been proven beyond a reasonable doubt.

Once a researcher has decided to use an implicit measure such as response latency, there is a further issue of how to characterize the central tendency of the person's responses to repeated exposures to the attitude object. Although many researchers have used mean scores, there are problems with this. For one, response latency distributions tend to be highly skewed in the direction of longer response latencies. This is because a large number of factors, such as blinking, distraction, and daydreaming, can increase response latencies. Consequently, mean scores are likely to present a distorted view of participants' actual central tendencies.

Several solutions to this problem have been proposed. The simplest solution is to use median, rather than mean, response latencies as the attitude measure. Another solution is to perform a reciprocal transformation of the data to reduce their skewness [$1/x$, $1/(x + 1)$ if any of the response latencies are less than 1 second, where x is the raw response latency]. A third solution is to perform a logarithmic transformation, which has the consequence of bringing the tail involving slower latencies closer to the center of the distribution. An additional recent solution involves a transformation of all of the latency data to z scores, and then mean or median z scores can be used as the attitude measure. In sum, there are several ways in which to deal with the problem of skewed latency distributions, and the choice of which method to use depends on considerations that are too numerous and complex to describe here.

10. DIRECT AND INDIRECT ATTITUDE MEASURES

Several expectancy–value attitude theories were proposed during the 1950s and 1960s. According to these theories, attitudes are a function of people's assumptions about the probability of various consequences arising from the performance of a behavior and evaluations of how good or bad those consequences are. Although the theories differ in the precise ways in which people are postulated to combine subjective probabilities and evaluations of consequences, they nevertheless have a common implication. Because attitudes are caused by a combination of beliefs about consequences and evaluations, this combination can be used as an indirect attitude measure. Consistent with expectancy–value theories, a large number of findings indicate that indirect and direct attitude measures are highly correlated, at least when the measures are created in accordance with the principle of correspondence. Consequently, many researchers have used indirect attitude measures, rather than direct ones, to predict other variables of interest, notably behavioral intentions and behaviors.

There has been some controversy about whether direct or indirect attitude measures are better predictors of behavioral intentions or behaviors. Although indirect measures have the advantage of specifying the subjective probabilities and evaluations of consequences that determine attitudes, there are both theoretical and empirical reasons to prefer direct attitude measures if one wishes to predict behavioral intentions or behaviors. The theoretical reason stems directly from the assumption that indirect measures assess variables (subjective probabilities and evaluations) that determine attitudes that, in turn, are a determinant of behavioral intentions or behaviors. According to this reasoning, attitudes are a more proximal cause of behavioral intentions or behaviors than are subjective probabilities and evaluations. Under the assumption that more proximal predictor variables work better than do less proximal ones, it follows that because direct attitude measures are assessing a variable (attitude) that is more proximal to behavioral intentions and behaviors than are indirect measures (which assess subjective probabilities and evaluations), direct measures should be a better predictor of behavioral intentions and behaviors than are indirect measures.

Numerous studies have provided tests of this theoretical reasoning. In the vast majority of cases, the reasoning has been supported; direct measures are generally

superior to indirect ones for predicting behavioral intentions and behaviors. An exception is when the attitude measure is not in accord with the principle of correspondence.

11. DISTINCTIONS THAT ARE IMPORTANT FOR ATTITUDE MEASUREMENT

With the problem of low attitude–behavior correlations having been solved by the principle of correspondence, recent researchers have changed their focus to three other problems that are now addressed. First, there is an important issue of whether attitudes are really an amalgamation of a cognitive (thinking) and affective (feeling) component. If there are separate cognitive and affective components, an implication is that each of these components should be measured separately, thereby improving the prediction of behaviors in a variety of applied domains. Second, there has been a great deal of controversy over whether attitudes should be measured separately from subjective norms. It was not until recently that this controversy was finally resolved. Finally, there are some attitudes that are not easily measured according to the principle of correspondence, and it might be necessary to make an adjustment. These issues are addressed in the following subsections.

11.1. The Separate Measurement of Cognition and Affect

Most researchers believe that there are both cognitive (thinking) and affective (feeling) components to attitudes. Three types of evidence support this belief. First, many researchers have used factor analysis to reduce a large number of attitude scales down to a smaller number of factors. In most cases, two factors result, with cognitive items loading on one factor and affective items loading on the other.

A second type of evidence comes from hierarchical regression analyses. Put simply, it is possible to consider the unique contribution of cognition to predicting general attitudes (or intentions) or to consider the unique contribution of affect. The results of these studies tend to indicate that, for any particular behavior, either cognition or affect will make a statistically significant unique contribution—a finding that should not occur regularly if attitude does not have these two components. An additional nicety of hierarchical regression analyses is

that the results are often consistent with researchers' intuitions. For example, cognition has been shown to be the more important contributor for what seems to be the cognitively controlled behavior of "studying over winter break," whereas affect has been shown to be the more important contributor for what seems to be the affectively controlled behavior of "smoking cigarettes."

Finally, Trafimow and Sheeran have made use of the associative hypothesis. They assumed that people have different types of beliefs about the consequences of behaviors and that some beliefs are more cognitive, whereas other beliefs are more affective. Given this, people are assumed to compare cognitive beliefs with other cognitive beliefs in the interest of forming the cognitive component of an attitude, and people are assumed to compare affective beliefs with other affective beliefs in the interest of forming the affective component of an attitude. But when people compare cognitive beliefs with other cognitive beliefs, or compare affective beliefs with other affective beliefs, they form associations; people form associations between cognitive beliefs and other cognitive beliefs, or between affective beliefs and other affective beliefs, but not between cognitive beliefs and affective beliefs. Consequently, when people are later asked to write down their beliefs about a behavior, writing a cognitive belief should cue the retrieval of another cognitive belief, whereas writing an affective belief should cue the retrieval of another affective one, thereby causing people's belief lists to be clustered by belief type. In contrast, if people do not distinguish between cognitive and affective beliefs (or between cognitive and affective components of attitudes), such clustering should not occur. In fact, such clustering is obtained, further supporting the distinction between cognitive and affective components of attitudes.

Empirical support for the distinction between cognitive and affective attitude components has led researchers to measure them separately to maximize the prediction of other variables such as behavioral intentions. Recent findings indicate that the prediction of behavioral intentions is significantly enhanced when the cognitive and affective attitude components are measured separately, and this finding has been replicated across a wide range of behavioral domains.

11.2. The Separate Measurement of Attitudes and Subjective Norms

Since the 1960s, researchers have assumed that attitudes and subjective norms are different causes of behaviors

(or, more often, behavioral intentions). Whereas attitudes have been assumed to be caused by beliefs about the personal consequences of performing a behavior, subjective norms have been assumed to be caused by beliefs about what important others think one should do. Because attitudes and subjective norms have been assumed to be different causes of behavioral intentions, it made sense to measure them separately to maximize the prediction of behavioral intentions. However, during the 1970s and 1980s, many researchers questioned the distinction between attitudes and subjective norms and argued that these were really different names for the same underlying idea. Obviously, if attitudes and subjective norms are different names for the same underlying construct, there is no reason to have distinct measures of each of them. Thus, there is an important measurement issue at stake in this controversy: Should attitudes be measured separately from subjective norms or not?

The argument against the distinction between attitudes and subjective norms was based on three issues. First, attitudes and subjective norms were often found to be highly correlated with each other, consistent with the notion that they are merely different names for the same underlying construct. Second, path analyses have sometimes indicated "crossover" effects, whereby attitudes and subjective norms affect each other. Third, philosophical arguments have been made that beliefs about consequences (i.e., behavioral beliefs) that are presumed to cause attitudes are not different from beliefs about the opinions of important others (i.e., normative beliefs) that are presumed to cause subjective norms. An example should make this issue clear. Suppose that someone has the behavioral belief that "my father will disagree if I eat chocolate" or the normative belief that "my father thinks I should not eat chocolate." Many researchers have argued that these two statements are just different ways of saying the same thing; therefore, if the cause of attitudes and subjective norms is the same, attitudes and subjective norms must also be the same.

Five kinds of findings have resolved most of the disagreement in favor of the distinction between attitudes and subjective norms. First, in general, behavioral beliefs have been found to be more highly correlated with attitudes than with subjective norms, and normative beliefs have been found to be more highly correlated with subjective norms than with attitudes—precisely what one would expect if behavioral beliefs cause attitudes and normative beliefs cause subjective norms. Second, although attitudes and subjective norms sometimes have been found to be highly correlated, the correlation often has been

found to be low or moderate. Third, attitudes and subjective norms have been manipulated experimentally, with different effects on behavioral intentions pertaining to different types of behaviors; manipulating attitudes has a larger effect on some behaviors, whereas manipulating subjective norms has a larger effect on others. Fourth, behavioral beliefs have been shown to be more strongly associated with other behavioral beliefs than with normative ones, and normative beliefs have been shown to be more strongly associated with other normative beliefs than with behavioral ones. Clearly, regardless of philosophical validity of the distinction between the two types of beliefs, people do make the distinction, contradicting the argument that attitudes and subjective norms have the same cause. Finally, there are individual differences between people in that some people are more under attitudinal control across a large number of behaviors, whereas others are more under normative control. If attitudes and subjective norms were the same thing, reliable individual differences in attitudinal or normative control should not be obtained. In sum, there is a great deal of evidence that attitudes and subjective norms are different from each other, and they should be measured separately.

12. CONSUMER APPLICATIONS

Attitude measurement is of more than just theoretical interest. Consumer and marketing researchers have applied the forgoing principles of attitude measurement in a variety of domains such as substance abuse among adolescents and tobacco use among college athletes as well as trying a new diet suppressant, getting mammograms, using condoms, drinking and driving, eliciting donations, purchasing environmentally friendly products, paying more for energy from renewable sources, purchasing food, voting, purchasing software, and even purchasing attitude research (by marketing directors). In these domains, as well as in many other domains, improved attitude measurement has resulted in an improvement in the prediction of behavioral intentions and behaviors from attitudes.

Improving the prediction of behavioral intentions and behaviors is not the only function of good attitude measures. It sometimes happens, although not very often, that an attitude does not do a good job of predicting a particular behavior. Before psychologists and consumer researchers had valid attitude measures, a low correlation between an attitude and a behavior was susceptible to at least two explanations. First, the

low correlation could show that the behavior is caused by something other than an attitude. Second, the low correlation could be due to an invalid attitude measure. An advantage of valid attitude measures is that they decrease the plausibility of the latter explanation and thereby increase the plausibility of the former one. Thus, when valid attitude measures nevertheless result in low correlations between attitudes and behaviors, researchers can be more confident in exploring other variables. A well-researched example of such an area is condom use. Although attitudes are capable of predicting condom use to some degree, other variables have also been shown to be good (perhaps better) predictors. Two of these are subjective norms and confidence that one knows whether others (e.g., one's sexual partner) think a condom should be used.

To illustrate the importance of valid attitude measures for consumer research regardless of whether the measures show that attitudes are a strong or weak predictor of behaviors, consider an example of a consumer researcher who wishes to increase sales of a particular product. Before investing money in an ad, the researcher needs to know what variables to address in the ad. If the behavior of buying the product is under attitudinal control for the population of interest, it makes sense for the ad to focus on variables that are likely to affect people's attitudes toward buying the product. Some of these variables might be beliefs about the product and affect. But what if the behavior is not under attitudinal control? In that case, there is little point in focusing an ad on variables designed to affect attitudes; for a behavior that is not under attitudinal control, there is no reason to believe that causing attitude change will increase sales. Consequently, the consumer researcher would be better served by creating an ad that focused on a different variable such as subjective norms. In general, having valid attitude measures increases the confidence that researchers can have in their data and provides a more solid basis for creating ads. Data showing that attitudes are good predictors of the behavior of concern provide a strong reason for creating an ad that is designed to affect attitudes; otherwise, the data provide a strong reason for creating an ad that focuses on other variables.

A further issue in attitude measurement, as applied to evaluating the effects of advertising, concerns the types of attitudes that should be measured. Consumer researchers have measured the effects of an ad on attitudes toward the ad, the product, or the brand. It has been rarer for researchers to measure the effects of an ad on attitudes toward the behavior the ad was designed to influence (e.g., buying the advertised product, voting for the advertised candidate). According to the principle of correspondence, if the reason for creating an ad is to influence behavior, it is precisely this last type of attitude that should be measured. Although there have been some demonstrations that an ad affects attitudes toward the behavior rather than merely attitudes toward the ad, the product, or the brand, more research is needed to establish the size of these effects. This is particularly so because the effects are likely to depend on a large number of variables such as the product domain, the type of ad, the type of use to which the product is put, and the frequency with which people are exposed to the ad.

One final example illustrates the importance of whether one measures attitudes toward the ad, the brand, the product, or the behavior of buying the product. Suppose that a consumer researcher for a company wishes to increase sales of a particular product. In addition, suppose that this researcher is evaluating an ad that has been shown to cause a change in people's attitudes toward the ad, the brand, or the product. It should be clear from the principle of correspondence that these attitudes, despite their seeming importance, are likely to not be particularly relevant to whether people will buy the product.

There may be cases where the consumer researcher's goal is something other than increasing sales of a particular product. Perhaps the goal is to increase sales of all of the products made by the company. In that case, it might be worthwhile to run an ad that focuses on increasing people's attitudes toward the brand. Even though such an ad might be unlikely to cause much of an increase in the sales of a particular product, a small increase in the sales of several products may justify the cost of the ad.

13. CONCLUSION

There has been a great deal of progress in how attitudes are measured. Current attitude measures are more reliable, more valid, and more correspondent to behaviors than were attitude measures in the past. In addition, more is known about some of the relevant issues that underlie attitude measurement. For example, recent research indicates that cognition and affect are both components of attitudes and should be measured separately. Recent research also indicates that attitudes are different from subjective norms and that the two variables should be measured separately. The consequence of such progress is that the prediction of behavioral intentions and behaviors from attitudes is

much greater than it has ever been before, and this has been demonstrated in a wide variety of domains.

See Also the Following Articles

Advertising and Culture ■ Advertising Psychology ■ Intentional Behavior ■ Measurement and Counseling

Further Reading

Ajzen, I. (1988). *Attitudes, personality and behavior.* Chicago: Dorsey.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice Hall.

Courneya, K. S. (1994). Predicting repeated behavior from intention: The issue of scale correspondence. *Journal of Applied Social Psychology*, *24*, 580–594.

Dovidio, J. F., Kawakami, K., & Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown, & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 175–197). Malden, MA: Blackwell.

Fazio, R. H. (1990a). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). San Diego: Academic Press.

Fazio, R. H. (1990b). A practical guide to the use of response latency in social psychological research. In C. Hendrick, & M. S. Clark (Eds.), *Review of personality and social psychology*, Vol. 2. *Research methods in personality and social psychology* (pp. 74–97). Newbury Park, CA: Sage.

Fishbein, M. (1980). Theory of reasoned action: Some applications and implications. In H. Howe, & M. Page (Eds.), *Nebraska Symposium on Motivation, 1979* (pp. 65–116). Lincoln: University of Nebraska Press.

Trafimow, D. (1998). Attitudinal and normative processes in health behavior. *Psychology and Health*, *13*, 307–317.

Trafimow, D., & Sheeran, P. (1998). Some tests of the distinction between cognitive and affective beliefs. *Journal of Experimental Social Psychology*, *34*, 378–397.

Attitudes

George Y. Bizer

Eastern Illinois University, Charleston, Illinois, USA

GLOSSARY

attitude A relatively enduring evaluation that a person holds about a target.

attitude change A process by which a person's evaluation toward a target is made more positive or negative.

attitude strength The extent to which an attitude is persistent over time, resists attempts to persuade, influences behavior, and influences cognition.

central route to persuasion The process of attitude change that relies on thoughtful processing of [the arguments within] a persuasive message.

Likert scale A tool for measuring attitudes in which respondents indicate to what extent they believe a series of items are characteristic of them regarding an attitude object.

peripheral route to persuasion The process of attitude change that does not rely on thoughtful processing of [the arguments within] a persuasive message.

semantic differential A tool for measuring attitudes in which respondents indicate how well a series of adjectives describes an attitude object.

tripartite model of attitudes Model suggesting that attitudes are derived from and/or influence affective, behavioral, and cognitive [components].

Attitudes are the relatively enduring evaluations that people hold toward all sorts of things in the world. Attitudes can be held toward people, objects, issues, and concepts. Researchers have devoted a great deal of effort to understanding attitudes because attitudes have been shown to be useful in predicting human behavior. Because of this interest, researchers have developed theories to explain where attitudes come from and tools with which to measure attitudes. Some research has focused on the distinctions between strong and weak attitudes, whereas other research has examined how attitudes can be changed. Because they are such an important part of everyday life, attitudes are one of psychology's most important constructs.

1. INTRODUCTION

A college student may like her psychology course but dislike her literature course. A consumer may prefer liquid laundry detergent over the powdered variety. And a city's populace may prefer the incumbent mayor over the challenger. Liking and disliking the objects in our world is a fundamental part of human nature. Psychologists use the term "attitude" to refer to these likes and dislikes that people hold—the relatively enduring evaluations that people hold toward all sorts of things. We can hold attitudes toward people (e.g., friends, celebrities), objects (e.g., automobiles, candy bars), political or social issues (e.g., capital punishment, immigration policy), and even abstract concepts (e.g., conservatism, democracy).

For years, researchers have spent a great deal of effort in attempting to better understand attitudes. One of the reasons why the concept has attracted so much attention is that attitudes serve an important role in our daily lives. The attitudes of college students influence which courses they decide to attend, the attitudes of consumers guide their decisions about which products to purchase, and the attitudes of a city's population affect voters' decisions about which candidate to elect. It is obvious that human behavior has a variety of causes, but understanding a person's attitudes is often a reliable way in which to predict the person's behavior. But attitudes also serve other roles. Sometimes, attitudes help us to gain rewards and avoid punishments. Other times, attitudes help us to define who we are as people; people may feel so strongly about an attitude that they define themselves in terms of that attitude (e.g., people may consider themselves to be "staunch Democrats" or "hardcore pro-lifers"). In addition, some research has even shown that holding negative attitudes toward other people may make us feel better about ourselves. Thus, because the attitudes we hold play important roles in many arenas, they are indeed a fundamental part of our everyday lives.

2. THE TRIPARTITE MODEL OF ATTITUDES

Where do attitudes come from? Research on the tripartite model of attitudes suggests that attitudes can be based on three different sources (for a visual representation of the tripartite model, see Fig. 1). One part of the tripartite model is the affective component. This affective basis of attitudes refers to the feelings, moods, and emotions that are associated with the attitude object. Consider a person who likes to exercise—a person who could be said to have a positive attitude toward exercise. This person's feelings of being "invigorated" or "distressed" whenever he or she exercises could serve as the affective basis of the attitude. Next, the behavioral basis refers to how the person has acted or would act in the future in regard to the target. If the person visits the gym to work out four times per week, this activity would serve as the behavioral basis. Finally, the cognitive basis refers to the overt thoughts or beliefs that the person has in regard to the target. If the person thinks that exercise is "healthy" and "worthwhile," such thoughts would serve as the cognitive component of the attitude.

In this example, the affective, behavioral, and cognitive components of the attitude all are positive, and this would suggest that the person will likely hold an overall positive attitude toward exercise, that is, the attitude object. However, the three components will not necessarily be consistent. A person may think that chocolate cake tastes wonderful (affective) and may eat chocolate cake often (behavioral) but may also think that chocolate cake is loaded with calories and fat (cognitive). In such cases, the overall attitude toward the object may be based more heavily on some components than on others.

It is important to note, however, that the components of attitudes can also be influenced by the attitude itself.
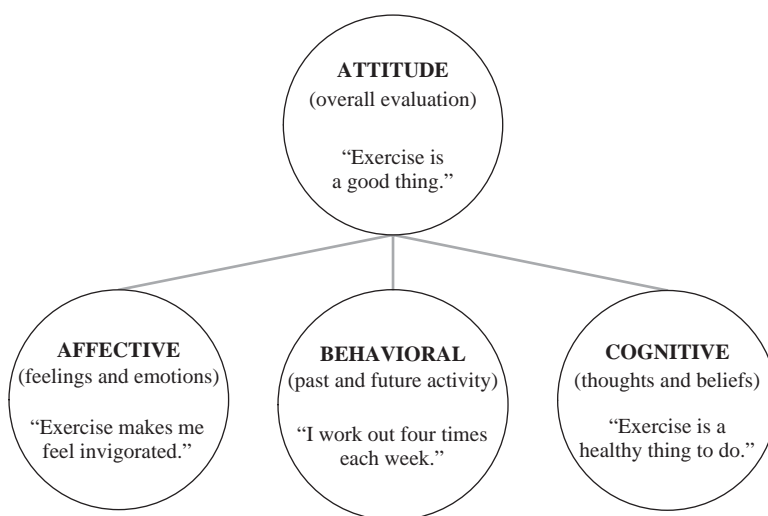

FIGURE 1   Visual representation of the tripartite model.

For example, in 1993, Eagly and Chaiken discussed how the attitude can also influence the three components. In other words, a person's affect, behaviors, and cognitions not only may guide and create the attitude but also may be driven or influenced by the attitude. As such, the three components of the tripartite model can be conceptualized in terms of both antecedents and consequences of their associated attitudes.

3. ATTITUDE MEASUREMENT

Because attitudes are so important, researchers have developed a variety of tools with which to measure them. Two of the most commonly used techniques are the Likert scale and the semantic differential. As shown in Fig. 2, participants reporting attitudes on a Likert scale are asked to what extent a variety of statements are characteristic of them. In a similar vein, as shown in Fig. 3, participants reporting attitudes on a semantic differential measure are asked to indicate how well a series of adjectives describe the attitude object. In either case, the researcher will typically provide a variety of items for the participant to complete. The scores from the items are combined to create an overall measure of the person's attitude toward the target.

There are several guidelines that researchers can follow in attempting to tap a person's attitude with as little error as possible. One guideline is to reverse code some of the items by framing them negatively as done in the third and fifth items in Figs. 2 and 3. Doing so

FIGURE 3 Example of a semantic differential scale.

serves two purposes. First, it reduces the likelihood that a person will simply check the leftmost or rightmost box down the entire instrument in attempts to save time or cognitive resources. A person is less likely to do so if he or she realizes that various items are indeed reverse coded. Similarly, if a person does simply check the leftmost or rightmost box down the entire instrument, the researcher may recognize this and treat the data accordingly. Also, some research suggests that providing respondents with five to seven options from which to select provides the most valid data. Scales with fewer than five options lack the precision necessary to measure attitudes accurately, whereas scales with more than seven options tend to suffer from additional noise without benefiting from any additional precision.

Likert scales and semantic differential scales are ubiquitous. Students typically use Likert scales when evaluating their professors at the end of a semester, and departments and universities use these data in part to make personnel decisions. Companies also use such techniques to measure customers' attitudes toward the products and/or services the companies offer. By learning which aspects of the companies customers like and dislike, managers can direct resources to improve deficient areas of their companies. In short, proper attitude measurement can gauge the likes and dislikes of a person or a group of people. Behavior of a person or a group of people can then be predicted based on the attitude measure.

4. ATTITUDE STRENGTH

Although attitudes typically predict behavior, this is not always the case. Some attitudes have tremendous

FIGURE 2 Example of a Likert scale.

influence over our lives, whereas others are largely inconsequential. For example, consider two people's attitudes toward environmental protection. The two people both may think that protecting the environment is a good thing—they both may hold positive attitudes toward environmentalism—but the strength of those attitudes may be quite different. One person's attitude may lead her to vote only for pro-environment candidates, drive a fuel-efficient car, and donate money to pro-environment organizations. On the other hand, the other person's attitude may have very little impact. He will occasionally recycle and has on occasion voted for pro-environment candidates. Why would two people who hold equivalent attitudes act so differently?

The difference lies in attitude strength. In 1995, Petty and Krosnick defined attitude strength as the extent to which an attitude persists over time, resists change when exposed to persuasive information, influences cognitive processes, and influences behavior (as in the preceding example). Although two people may hold what appear to be equivalent attitudes—perhaps as observed by identical scores from Likert scales—the attitudes might not be functionally equivalent. Researchers have identified a wide variety of ways in which attitudes can become stronger, including thinking about attitudes, learning about the attitude object, or expressing attitudes repeatedly.

Attitude strength can be measured by observing the four defining features directly: how long attitudes persist over time, how well attitudes resist persuasive attempts, the extent to which attitudes predict cognitive processes, and the extent to which attitudes predict behavioral processes. But there is a wide variety of other features that are associated with these defining features. Some researchers have investigated the certainty with which attitudes are held: when a person feels certain of an attitude, that attitude is likely to show the four defining features of attitude strength. In a similar vein, researchers have shown that accessible attitudes, attitudes based on much knowledge, and attitudes on issues that people find to be personally important also tend to manifest the four defining features of attitude strength.

Therefore, many researchers measure not only attitudes but also the strength of those attitudes. For example, consider an incumbent governor's campaign manager determining in which counties to invest scarce advertising funds. Although two counties may show roughly equivalent positive attitudes toward the incumbent, such a report tells only part of the story. Based on this information alone, the manager might decide to split the advertising dollars across both counties equally. However, it is possible that the citizens of County A may be, on average, highly certain of their positive attitudes, whereas the citizens of County B may be, on average, not certain at all. In such a case, one would expect that the positive attitudes would predict vote choice better among the residents of County A than among those of County B. Therefore, the manager would benefit more from investing advertising in the county with the weaker attitudes (as measured by certainty) than in the county with the stronger attitudes. Thus, measuring attitudes alone is only part of the story. One must also measure the strength of those attitudes to predict the behavioral and cognitive consequences of attitudes most effectively.

5. ATTITUDE CHANGE

Attitudes are thought to be relatively enduring evaluations. However, there is a wealth of research on how attitudes can and do change in the face of a persuasive communication. One of the most widely studied theories of persuasion is Petty and Cacioppo's elaboration likelihood model. In their 1986 article, these researchers posited that there are two routes through which persuasion can take place. When people have the motivation and the ability to carefully process and think about the arguments within a counterattitudinal message, they will use the central route to persuasion. When people process messages centrally, they evaluate the arguments within the message—the logical factual information that is directly relevant to the attitude object. People will then change their attitudes to the extent that their cognitive responses—the thoughts they have about the arguments while reading the message—are positive. However, when people lack either the motivation or the ability to carefully process and think about the arguments within a counterattitudinal message, they will use the peripheral route to persuasion. When people process a message peripherally, instead of elaborating on the arguments within the message, they respond to cues that exist within the message—images or heuristics not directly relevant to the message that can persuade without requiring much thought.

For example, consider an advertisement for a new car that appears in a general-interest magazine. At the top of the advertisement is the name of the car, and at the bottom of the page are three paragraphs detailing the car's resale value, horsepower, gas mileage, and safety record. In the background is an image of the car being driven across a sun-drenched beach with beautiful palm trees swaying in the background. A reader who is in the market for a new automobile may have the motivation to

use the central route to persuasion when processing the advertisement. This person will carefully read the arguments at the bottom of the page and, assuming that the reader's cognitive responses are positive, will end up with a more positive attitude toward the car. However, consider another reader who does not plan to purchase a car for years and therefore has no motivation to use the central route. This person may instead simply observe the beautiful palm trees and the beach and therefore simply create an association between the positive imagery and the car. Thus, the second person may also end up with a more positive attitude toward the car but will do so through the peripheral route instead of the central route.

If both the central and peripheral routes can lead to the same effect (e.g., attitude change), why does it matter which route to persuasion people take? Perhaps the most important reason is that attitudes resulting from central route persuasion tend to be stronger than attitudes resulting from peripheral route persuasion. That is, attitudes formed from central route processing tend to be more persistent over time, more resistant to future persuasion, and more impactful on behavior and cognitive processes than do attitudes formed from peripheral route processing. So, from this standpoint, it would seem better to engage an audience in central route processing due to the enhanced strength that would result. However, peripheral route processing may be the best option when (a) the audience does not have the ability or motivation to use the central route or (b) there are very few compelling arguments that can be presented in a message in the first place. In short, persuasion practitioners must balance the costs and benefits when determining whether to use primary central route or peripheral route persuasive appeals.

6. CONCLUSION

Attitudes, as the representations of what we like and dislike, play an important role in our daily lives. From important decisions such as choosing one presidential candidate over another to more mundane decisions such as which flavor of ice cream to purchase, attitudes play a fundamental role in all sorts of behavioral and cognitive processes. Because they are such an important part of our lives, they have all sorts of important applications. Political scientists study attitudes in attempting to predict voters' choices, consumer psychologists study attitudes in attempting to understand consumers' purchasing decisions, and health promotion researchers study attitudes in attempting to understand the healthy and unhealthy behaviors in which people engage. In short, because they play a fundamental role in explaining behavioral and mental processes, attitudes are one of psychology's most important constructs.

See Also the Following Articles

Attitude Measurement ■ Intentional Behavior ■ Measurement and Counseling

Further Reading

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 269–322). New York: McGraw–Hill.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*, 163–204.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205.

Petty, R. E., & Krosnick, J. A. (Eds.). (1995). *Attitude strength: Antecedents and consequences*. Hillsdale, NJ: Lawrence Erlbaum.

Zimbardo, P. G., & Lieppe, M. R. (1991). *The psychology of attitude change and social influence*. New York: McGraw–Hill.

Authoritarianism

Klaus Boehnke

International University Bremen, Bremen, Germany

Andreas Hadjar

Chemnitz University of Technology, Chemnitz, Germany

GLOSSARY

attitude Persistent mental state of readiness to react to a certain object or class of objects in a certain way; learned disposition to behave in a consistent way toward a given class of objects.

disposition Relatively long-lasting emotionally rooted attitude.

ethnocentrism Tendency to divide the social world into groups with which one identifies and to which one submits (in-groups) and groups of outsiders (out-groups) to which one is hostile; characterized by glorification of the in-group and defamation of the out-group.

personality structure Unit that underlies individual ways of behaving and gives consistency to otherwise contradictory-seeming mannerisms and behavioral or attitudinal specialties.

personality syndrome Acquired pattern of personality characteristics bearing resemblance to the personality structures of others who shared similar experiences and problems and made similar adaptations.

xenophobia Abnormal fear of strangers.

Authoritarianism is a psychological concept that assumes the existence of a more or less unconscious general orientation of individuals toward being antidemocratic, prejudiced, and fascist. The so-called "authoritarian personality" is characterized by two ambivalent tendencies: subordination to authorities and the desire to become an authority himself or herself. Authoritarianism as a personality syndrome is predominantly measured by using the so-called F-scale and later revisions of that instrument.

1. AUTHORITARIANISM: THE CONCEPT

There is no concept in the social sciences that is more closely connected to the events in Germany and throughout the world in the 1930s and 1940s than the authoritarianism concept. For decades, this approach served as the main explanation of fascism and anti-democratic thoughts or action in psychology.

1.1. Historical Background

Precursors of the authoritarianism concept can be traced back to the early and mid-1930s. At that time Marxist psychoanalyst Wilhelm Reich studied voting behavior of Germans between 1929 and 1933 to find reasons for the increase in the number of voters for the fascist NSDAP party from less than 1 million to approximately 17 million voters during that period. Within his framework of thinking, he concluded that ideologies of subordination to authority were

internalized by subordinate individuals and eventually became a stable personality structure. Some years later, researchers of the so-called Frankfurt School, such as Erich Fromm, Theodor Adorno, Max Horkheimer, and Herbert Marcuse, carried out a study on "Authority and Family" and systematized the idea of an authoritarian personality structure that is strongly determined by the societal context. In the 1940s, members of the Frankfurt School, who had to leave Germany due to their Jewish and Marxist background, engaged in a much larger research project—partly supported by the American Jewish Committee—in their American exile. Influenced by their experience of racism and inhumanity in fascist Germany, the group around Adorno set out to uncover psychological forces that promote fascist and antidemocratic attitudes and to find possible ways to fight fascist tendencies in society. They used quantitative and qualitative research methods (e.g., questionnaire-based surveys and qualitative semistructured interviews). Individuals surveyed in the U.S.-based studies were mainly students and members of the American middle class.

1.2. Concept of the Authoritarian Personality

The main thesis of Adorno was that the political, economic, and social attitudes of an individual form a coherent structure of thought, which is an expression of an underlying, hidden personality structure. In accordance with Freudian psychodynamic theory, this personality structure is assumed to have developed during early childhood. Authoritarian relations within the family, an authoritarian parental style, and a general lack of love and warmth determine it. Antidemocratic individuals often experience feelings of hatred against authorities—namely their father as the family's bread winner—during their childhood. Because they never could express this hatred in any way against their father or their parents, they direct such feelings of hatred toward convenient scapegoats—inferiors and social minorities (e.g., immigrants/foreigners, homosexuals, and the handicapped)—later in life. Main characteristics of authoritarian personalities include a preoccupation with superiority and one-upmanship, an intolerance of behavior and objects that are different, a tendency to classify all things into black and white/right and wrong, a strict adherence to received views and prevailing social trends, and superficial respect for authority figures. Adorno and collaborators developed the so-called F-scale (fascism scale) to measure such an antidemocratic personality structure (Fig. 1).

The scale is constructed to measure a syndrome that comprises nine dimensions or subscales not always present simultaneously:

1. Conventionalism: rigid adherence to the conventional values of the middle class
2. Authoritarian submission: an uncritical and submissive attitude toward idealized in-group authorities
3. Authoritarian aggression: a tendency to look out for and to condemn, reject, and punish people who violate conventional values
4. Anti-intraception: rejection of all inwardness, of the subjective, the imaginative, the tender-minded, and of self-criticism
5. Superstition and stereotypicality: the belief in mystical determinants of the individual's fate, a disposition to think in rigid categories
6. Power and toughness: a preoccupation with the dimension of dominance–submission, strong–weak, leader–follower; identification with power figures; overemphasis on the conventionalized attributes of the ego; exaggerated assertion of strength and toughness
7. Destructiveness and cynicism: a generalized hostility and vilification of the human
8. Projectivity: to project unconscious emotional impulses onto the outside world and to believe that wild, evil, and dangerous things go on in the world
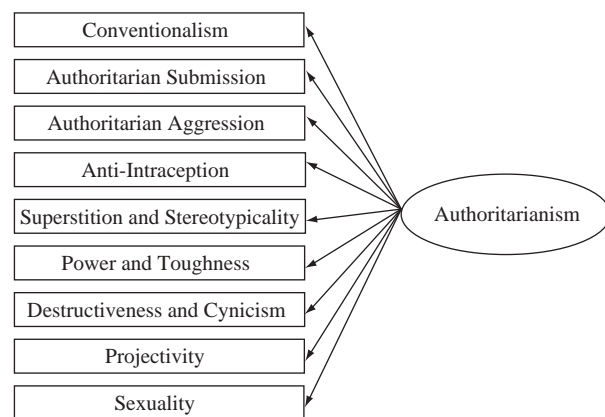9. Sexuality: an exaggerated concern with sexual matters

FIGURE 1   The authoritarianism construct.

1.3. Criticism

The accomplishment of Adorno and collaborators was to create an insight into the phenomenon of an authoritarian personality. However, there are methodological and empirical limitations. One of the foremost points of criticism is that the theory attempts to explain societal processes (Germany's turn to fascism) on the basis of knowledge about individual personality development. Another major criticism is that the authoritarianism concept and the F-scale questions appear highly ideological. This criticism is repudiated by a number of psychologists, including Gerda Lederer, who points out that F-scale items are related to moral and individual values, human relations, the self, family, and sexuality and are not at all ideological. At least the popular reception of the concept, however, is highly politicized. Particularly in the 1950s, it was used to characterize not only fascist but also communist movements. Yet other points of criticism are concerned with methodological problems of Adorno's empirical work, indicating, for instance, a sampling problem. Adorno generalized from a nonprobability sample that consisted of white middle-class Americans. Also, several studies showed problems with the construct validity of the F-scale, i.e., they doubt that the F-scale measures exactly what it intends to measure. Furthermore, semistructured interviews conducted with individuals in order to gain information on their parents' educational style were criticized as being more of a memory exercise than a pool of information about parental authoritarianism. Some studies replicating Adorno's work found no evidence to support the central findings, particularly on the impact of authoritarian socialization. One study even reports evidence for a positive relationship between authoritarianism and xenophilia. However, taking into account the multitude of existing studies, one can see the existence of the phenomenon authoritarian personality as confirmed.

1.4. Contemporary Developments of the Authoritarianism Concept

The initiator of the new research stream on authoritarianism is Robert Altemeyer. His approach leaves the theoretical basis of the Freudian psychodynamic theory behind and focuses on learning processes highlighting agents of socialization, such as family and peer group, and the context of society. In his work the so-called right-wing authoritarianism is influenced by situation or environment rather than just personality development. Altemeyer reduces the dimensions of authoritarianism to three: authoritarian submission, authoritarian aggression, and conventionalism. In the 1980s, many social scientists attempted to develop new instruments to measure the authoritarian personality. Others tried to explore by qualitative research methods the conditions that lead to such a personality structure.

1.5. Similar Contemporary Concepts

There are two contemporary competitors to the authoritarianism concept. During the past decade, Sidanius and Pratto studied what they call the social dominance orientation (SDO). SDO refers to fundamental values to accept social hierarchies. These values form a personality structure in the Freudian sense and determine so-called legitimizing myths, which are moral and intellectual justifications of individual or institutional discrimination. Whereas SDO still recognizes psychodynamic assumptions, the concept of hierarchic self-interest (HSI) is based on a paradigm of lifelong socialization in a market-oriented economy. HSI is an expression of the individual effort to perform "better than others" in a hierarchically structured society, an internalized elbow mentality, and it appears as a syndrome of values that comprises three core dimensions (success orientation, competitiveness, and individualism), as Hagan and collaborators showed in their youth studies. Since HSI is a syndrome and second-order construct (see Fig. 2 for an example from empirical research), it is not ultimately determined but can
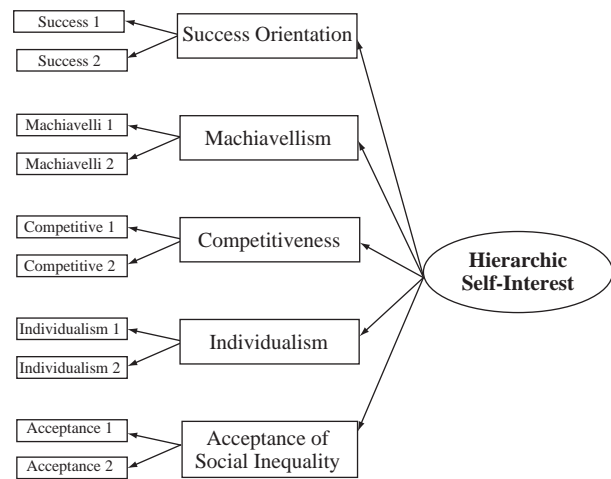

FIGURE 2 Hierarchic self-interest (structural model).

be modified by integrating other first-order factors (e.g., machiavellism and acceptance of social inequality). HSI turned out to be a stable predictor of xenophobia. This finding is in a way paradoxical because market-oriented HSI values apparently lead—under the condition of increasingly limited resources—to a tendency of ethnocentric market restriction. Roots of HSI are seen in modes of workforce participation of men and women and parental styles so that high HSI values may be interpreted as expressions of familial or socioeconomic disintegration.

2. CORRELATES OF AUTHORITARIANISM

2.1. Antecedents

2.1.1. Class and Socioeconomic Status

Empirical evidence suggests that authoritarianism is found among the lower classes to a higher degree than among the middle classes. Lipset assumes in his working class authoritarianism thesis that lower classes tend to have an authoritarian view on politics; therefore, they are more likely to support extremist movements that promise fast and simple solutions to social problems and that blame inferior scapegoats for the problems. This authoritarian view results from low education, the degree of isolation of the class, economic and psychological uncertainty, and the particular family life in lower classes. Depending on the exact kind of authoritarianism scale that was used and how class was measured, results of empirical studies support this thesis, particularly the influence of education on authoritarianism. A higher level of education seems to be the decisive characteristic of nonauthoritarian people.

2.1.2. Family Socialization

Adorno *et al.* emphasize the role of the family for the development of the authoritarian personality. Other scientists studied the mechanisms that lie behind the relationship of socialization and authoritarianism in more detail. They suggested that the transmission of authoritarian attitudes from the parents onto the children depends heavily on the strength of social and emotional ties between parents and children. Central to an authoritarian socialization are parental styles that include violence and do not recognize the

needs of the children. Authoritarian families usually are ruled by the father and do not allow children to participate in family decisions.

2.1.3. Cross-Cultural Differences

In dealing with the concept of authoritarianism, it must be discussed whether this concept applies only to the Western world or whether it is universal. Based on data from more than 100 countries, Meloen shows that there is a strong relationship between culture, attitudes, and politics. State authoritarianism is highly related to authoritarian attitudes among the citizens. These attitudes seem to result from a culture that is based on hierarchies and a traditional family structure. Comparisons of authoritarianism among adolescents in East and West Germany after the fall of the Berlin Wall show very clearly that different socialization cultures lead to different degrees of authoritarianism: East German adolescents agreed with authoritarian statements more than West Germans did. However, scales to measure authoritarianism cannot be used throughout the world in the same way and the same composition. Words and contexts are interpreted differently in different cultures. Studies in Japan, for example, showed that authoritarian structures and authoritarian attitudes do exist but cannot be measured in terms of American authoritarianism instruments.

2.2. Consequences

The best studied and most powerful relation is the one obtained for authoritarianism and ethnocentrism. Ethnocentrism is an attitude characterized by the glorification of one's own group (in-group) and the defamation and discrimination of other groups (out-group). Xenophobia, racism, and nationalism are other orientations similar to ethnocentrism. An authoritarian personality tends to be more ethnocentric. In certain circumstances (situational factors), the just mentioned attitudes can lead to discrimination, violence, and hate crimes.

Authoritarianism is also connected to other attitudes, namely sexism, anti-Semitism, cognitive rigidity (dogmatism), and political and economic conservatism. Authoritarian personalities tend to devalue women, homosexuals, and Jews. They strongly support the market economic system and back conservative political movements. Authoritarianism also may explain oppression, chauvinism, and negative put-downs on

the individual level and the level of society. Many of the empirical findings on consequences of authoritarianism may, however, rightly be accused of bearing an element of tautology because almost all consequences addressed could also be seen as elements of the authoritarian syndrome per se.

3. TREATMENT APPROACHES

From knowledge about explanatory factors of authoritarianism, particularly socialization, a list of treatment measures can be derived. The founders of the Frankfurt School suggested enlightenment and an education fostering responsibility. That means making the mechanisms of society transparent to a broad audience. The psychologist Kurt Lewin, who conducted several experiments on group conflicts in small groups, pointed out that democratically acting group leaders are necessary to prevent authoritarianism. Group leaders as well as all other group members must be enabled to play their role in the democratic process. Authoritarian leadership styles must be replaced by consultative and participative leadership styles. To support this process of democratization, the democratic leader has to leave space for self-determination with clearly defined and reasoned borders to the group members rather than predetermining a more or less restricted space with rigid borders. Democratic skills cannot be passed on in a simple way; they must be learned and interactively developed in every generation, in all institutions of society, including the family, the school, and the workplace. Whereas a Freudian approach to authoritarianism would suggest that there is no treatment for authoritarian personalities in adulthood, from the perspective of lifelong socialization a cure seems possible. To push back authoritarian tendencies in daily life settings, people's integration into groups (working groups, peer groups, and families) must be supported. People must be given participation opportunities in group decisions; rules, actions, and decisions shall be explained to them in a detailed and reasoned way; they shall be encouraged to look at themselves and others from different perspectives; and every kind of rigid discipline and physical or mental violence must be absent in dealing with people who tend to be authoritarian.

See Also the Following Articles

Family and Culture ■ Industrial/Organizational Psychology across Cultures ■ Interpersonal Behavior and Culture ■ Power, Authority, and Leadership ■ Stereotypes

Further Reading

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper.

Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism*. San Francisco: Jossey-Bass.

Brown, R. (1995). *Prejudice. Its social psychology*. Oxford: Blackwell.

Christie, R., & Jahoda, M. (1954). *Studies in the scope and method of the authoritarian personality*. Glencoe, IL: Free Press.

Hadjar, A. (2004). *Ellenbogenmentalität und Fremdenfeindlichkeit bei Jugendlichen. Die Rolle des Hierarchischen Selbstinteresses* [*Elbow mentality and xenophobia among adolescents. The role of hierarchic self-interest*]. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Hagan, J., Hefler, G., Classen, G., Boehnke, K., & Merkens, H. (1998). Subterranean sources of subcultural delinquency beyond the American Dream. *Criminology*, *36*.

Lewin, K. (1948). *Resolving social conflicts*. New York: Harper & Brothers.

Martin, J. L. (2000). The authoritarian personality, 50 years later: What lessons are there for political psychology. *Political Psychology*, *22*.

Meleon, J. D., Hagendoorn, L., Raaijmakers, Q., & Visser, L. (1988). Authoritarianism and the revival of political racism: Reassessments in The Netherlands of the reliability and validity of the concept of authoritarianism by Adorno *et al. Political Psychology*, *9*.

Rippl, S., Seipel, C., & Kindervater, A. (2000). *Autoritarismus. Kontroversen und Ansätze der aktuellen Autoritarismusforschung* [*Authoritarianism: Controversies and approaches of contemporary research on authoritarianism*]. Opladen: Leske & Budrich.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Aviation

Evan A. Byrne

National Transportation Safety Board, Washington, D.C., USA

GLOSSARY

aviation psychology Scientific discipline focused on understanding and improving human performance in aviation.

crew resource management A collection of procedures, techniques, and communications strategies that fosters the use of all available resources (including people and available technology) to provide optimal levels of safety and mitigate the consequences of error.

cultural factors Procedural, communications, equipment, and design-related issues that are unique to an organization or a society.

electroencephalogram (EEG) Measurement of aggregate brain electrical activity recorded at the scalp.

FAA Abbreviation for Federal Aviation Administration.

NASA Abbreviation for National Aeronautics and Space Administration.

National Airspace System The complex system of people, equipment, and procedures that guides how aircraft are operated in the United States.

NTSB Abbreviation for National Transportation Safety Board.

situation awareness How people keep track of critical features in their environment, such as their position in space and time, and the current and future status of systems.

Aviation psychology is a simple name for a deceptively broad multidisciplinary applied domain. Its goal is plain: understanding and improving human performance in aviation. As a formal discipline, it draws on other established areas of psychology and engineering. For example, within the discipline of psychology, aspects of applied, experimental, social, clinical, industrial, organizational, and engineering specializations have been incorporated into aviation psychology.

1. HISTORY

The origins of aviation psychology are nearly as old as powered flight. Whether it played a role in the selection of the Wright *Flyer*'s first pilot on the sands of Kitty Hawk, North Carolina, a century ago is debatable. Lindbergh's self-account of his 33-hour transatlantic crossing in the *Spirit of St. Louis* in 1927 is among the earliest accounts of the effects of sustained wakefulness on pilot performance. However, aviation psychology is generally described as having its formal beginning coincident with the increased military use of aviation in the two world wars. During those times, there was a fundamental need to rapidly select and adequately train the pilots required to fly the fighters, bombers, and transport airplanes for the war efforts. As aviation became increasingly important to worldwide commerce, and with the advent of the space program, aviation psychology as a discipline expanded during the postwar years. This expansion paralleled technological

advances that increased the complexity of some facets of aviation and the continuing need to understand human performance capabilities and limitations in these domains. Several detailed recent reviews of the domain exist that expand on the history and basic applications in this area.

2. SCOPE AND EXTENT OF THE DOMAIN

Enhancing pilot performance is a major focus of aviation psychology; however, its reach in aviation extends far beyond pilots. For example, early efforts of aviation psychologists included enhancing the training of aircraft gunners and selection of aircraft mechanics. Aviation psychology focuses on understanding the operator's basic limitations and needs. Basic needs for optimal performance typically include adequate selection and training, sufficient real-time information, appropriate human–machine interfaces and equipment, and compensated environmental conditions. The scope of aviation psychology has generally kept pace with the rapid expansion of aviation itself since the advent of powered flight.

The National Airspace System is described on the Federal Aviation Administration's (FAA) Web site as "a complex collection of systems, procedures, facilities, aircraft, and ... people [that] represents the overall environment for the safe operation of aircraft." With its focus on the people and equipment involved in this complex system, aviation psychology is broad in scope but singular in purpose: to understand and promote optimal human performance. Aviation psychologists have contributed to our knowledge of human performance in both civilian and military applications. They have examined and sought ways in which to improve the performance of pilots, flight attendants, mechanics, air traffic controllers, and ground support personnel. It is clear that aviation psychologists now also have an important contribution to make in the study of human performance factors in civil aviation security, for example, in selecting and training security screening personnel and in identifying methods to augment and support human search capabilities during repetitive inspection tasks.

A peer-reviewed journal, the *International Journal of Aviation Psychology,* serves as a central focus for the dissemination of research in this area, and a survey of its contents can quickly provide a comprehensive overview of this area. Other professional journals, such as *Human Factors* (published by the Human Factors and Ergonomics Society) and *Aviation, Space, and Environmental Medicine* (published by the Aerospace Medical Association), also contain work by aviation psychologists. In the latter case, there are some who may be more appropriately titled "aerospace psychologists" given that their focus involves understanding human performance in space-related applications such as low-*g* environments and crew coordination issues during long-duration space flights.

3. SOURCES OF BASIC PRINCIPLES AND THEORETICAL APPROACHES

Aviation psychology draws on basic principles, findings, and tenets from other specialized domains of psychology, sociology, physiology, and human factors. For example, knowledge of sensory processes, memory limitations, psychomotor skills, and cognitive functioning translates directly to the aviation environment and can be used to optimize the design of instruments, procedures, and warning devices. Other areas of experimental and cognitive psychology that provide insight into attention and decision making also contribute to aviation psychology in areas such as evaluating instrument scan behavior, understanding visual detection capabilities for pilots (or inspectors detecting defects in manufacturing/maintenance inspection processes), and aeronautical decision-making skills. Finally, knowledge of basic human performance capabilities under exposure to noise, vibration, motion, and altitude facilitates aviation psychologists' understanding of human capabilities and limitations.

Methods derived from industrial, organizational, and educational psychology can aid in the selection and training of aviation personnel. Components from cognitive psychology, social psychology, and sociology can contribute to the study of crew coordination and cultural factors in aviation. For example, cultural factors in aviation are not just associated with design stereotypes for the movement of controls and switches. They also include crew coordination issues that can vary across cultures, include variations in the acceptance of automated technology, and can even affect the coordination of operating practices when airlines merge.

Physiological psychology has contributed to the measurement of workload and assessment of fatigue in aviation environments. Although few operational

examples of the former exist, the use of physiological measures to infer the state of operators has taken place in basic and applied aviation research environments. These include traditional measures such as heart rate and electroencephalogram (EEG). In addition, eye movements have been used to study operator scan patterns for complex displays or while performing aviation-related tasks.

There are few, if any, theoretical approaches unique to aviation psychology. Instead, basic theories generated elsewhere are applicable to basic and applied questions in aviation. Some theoretical or worldview approaches are especially salient as applied to the aviation environment. For example, information processing theories that outline and predict how operators process information from the environment and manage conditions of increasing workload are useful in evaluating aviation work environments to ensure that workload and capabilities are not exceeded. Also, during recent years, significant effort has been focused on understanding and establishing theoretical and methodological approaches to the study of situation awareness, that is, how operators learn about and keep track of critical features in their environment, such as their position in space and time, and the current and future status of systems. Theoretical models for the acquisition and maintenance of situation awareness by personnel involved in aviation tasks have been developed, although acceptance of their utility is not universal.

Aviation psychology presents a unique application and research environment. However, with few exceptions (e.g., operations in extreme environments, operations under conditions of g loading, specific examination of flight- or domain-specific tasks such as the visual processes involved in landing), knowledge, methods, and concepts derived elsewhere in psychology and engineering can readily translate to applications and research questions in this area. For additional information on the basic tenets, principles, and theoretical approaches used by aviation psychologists, there exist recent compendiums of work outlining these issues from both historical and contemporary perspectives.

4. APPLICATIONS AND ONGOING ISSUES

Aviation psychology is an incremental and evolving discipline that is generally responsive to the dynamic challenges it faces as aviation itself continues to evolve.

Following are some of the areas of inquiry in which aviation psychologists are engaged.

4.1. Error Management

Studies and training methods for crew coordination have evolved from attending static personality and "charm school" sessions to delivering strategies for the prompt identification and timely recovery from human errors that are inevitable. Work continues on determining optimal methods for enhancing crew coordination and error management, on developing instructional methods for distributing these error management concepts through technical training, and on reinforcing during routine operations through robust procedures.

4.2. Decision Making and High-Stress Situations

Research continues on understanding how personnel in the aviation environment arrive at decisions under conditions of stress, often having limited information. Some of these efforts are targeted to developing methods to enhance pilot decision making and judgment. Other efforts examine factors that affect human performance in stressful situations such as passengers evacuating an airplane under an emergency situation.

4.3. Fatigue Countermeasures

Substantial work in the area of fatigue has been performed by aviation psychologists at the National Aeronautical and Space Administration (NASA) Ames Research Center, and work continues to better understand the effects of fatigue on aviation operations. Research conducted by the FAA has looked at shiftwork and fatigue in air traffic control operations. Aviation psychologists working for the military have examined scheduling practices and pharmacological countermeasures to mitigate the effects of fatigue in military aviation.

4.4. Automation and Human–Machine Interfaces

The study of human interaction with automation remains a dominant focus for many aviation psychologists. The use of automated aids is increasing in both air carrier and general aviation, and this trend is expected to continue with the refinement of more complex aircraft

system status displays, horizontal and vertical navigation displays, airport situation awareness displays, highway-in-the-sky flight path aids, and the like. Automated aids are also increasingly being applied to other aviation operations such as air traffic control and maintenance.

4.5. New Training Technologies

Simulators have found widespread use in the training of air carrier flight crews. Methods for enhancing the training environment using these tools continue to be evaluated. The use of computer-based training and part task simulators is also spreading to other areas of aviation, including general aviation and maintenance technicians. Ensuring that the transfer of training effects is positive and predictable will continue to be a focus for aviation psychologists.

4.6. Accident Reduction

Methods to study human error in accidents continue to be examined to identify broad causal factors common to multiple accidents for which targeted intervention strategies could be applied. Strategies for this work include the use of surveys, the use of simulator studies, and review of accident data.

4.7. Maintenance and Other Aviation Applications

Some areas that have received relatively less attention in the past, as compared with flight deck and pilot issues, are receiving increasing attention by aviation psychologists. These include maintenance and ground support personnel, and recent accidents have shown that human error in these areas can also yield catastrophic outcomes. Strategies for increasing performance on the flight deck, such as crew resource management, are being translated to optimize performance by operators in these areas. Similarly, additional focus is being placed on the information needs of maintenance personnel to complement work done previously to examine environmental and ergonomics needs.

5. EDUCATION AND TRAINING

Because of the multidisciplinary characteristics of aviation psychology, there is no single training or education path to become an aviation psychologist (although for decades some institutions, such as the University of Illinois and Ohio State University, have produced professionals who made significant contributions in this area). Like the applied domain of human factors, successful performance in aviation psychology requires a professional who has developed a strong grounding in basic psychological principles, experimental methodology, and basic familiarity with engineering concepts. Creativity and the ability to adapt are essential traits that facilitate interaction with experts from the many facets of aviation, and formulation of strategies to cope with the short-cycle time periods (from question to required answer) characteristic of many issues aviation psychologists may be called on to study.

Equally important to becoming an effective aviation psychologist is a strong interest in, some observational experience in, or (optimally) direct experience in an operational facet of aviation (e.g., holding a pilot certificate). This experience will also help to facilitate communication with subject matter experts and provide the aviation psychologist with practical knowledge that will augment his or her professional training in a chosen area of psychology or in an allied discipline. Aviation has many regulatory and procedural boundaries that have been established by regulators, manufacturers, companies, and labor groups. These areas, whether they involve hours of service regulations for pilots or the maintenance and inspection routine, should be mastered to the extent possible for an aviation psychologist to contribute optimally.

6. OCCUPATIONS AND EMPLOYMENT

Aviation psychologists can be found employed by organizations in and on the periphery of the aerospace industry. For example, they are employed by agencies such as NASA, the FAA, and the National Transportation Safety Board (NTSB) in the federal government and are involved in basic research, certification, design, oversight, and safety investigations. In addition, most branches of the U.S. Department of Defense (DoD) have similar areas in which aviation psychologists are employed to facilitate military aviation. In the private sector, aviation psychologists are employed by manufacturers, airlines, and private research institutions. Academia is another area where aviation psychologists can be found working on basic and applied research questions, often funded by DoD, NASA, and/or FAA grants and projects.

See Also the Following Articles

Accidents in Transportation ■ Engineering Psychology ■ Transportation Systems, Overview

Further Reading

Billings, C. (1997). *Aviation automation: The search for a human-centered approach.* Mahwah, NJ: Lawrence Erlbaum.

Caldwell, J. A., & Gilreath, S. R. (2002). A survey of aircrew fatigue in a sample of U.S. Army aviation personnel. *Aviation, Space, and Environmental Medicine*, 73, 472–480.

Dismukes, R. K. (1994). Aviation human factors research in U.S. universities: Potential contributions to national needs. *International Journal of Aviation Psychology*, 4, 315–331.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32–64.

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors*, 27, 149–157.

Garland, D. J., Wise, J. A., & Hopkins, V. D. (Eds.). (1999). *Handbook of aviation human factors.* Mahwah, NJ: Lawrence Erlbaum.

Guilford, J. P. (1948). Some lessons from aviation psychology. *American Psychologist*, 3, 3–11.

Helmreich, R. L., & Merritt, A. C. (1998). *Culture at work in aviation and medicine: National, organizational, and professional influences.* Aldershot, UK: Ashgate.

Koonce, J. M. (1984). A brief history of aviation psychology. *Human Factors*, 26, 499–508.

O'Hare, D., & Chalmers, D. (1999). The incidence of incidents: A nationwide study of flight experience and exposure to accidents and incidents. *International Journal of Aviation Psychology*, 9(1), 1–18.

O'Hare, D., & Lawrence, B. (2000). The shape of aviation psychology: A review of articles published in the first 5 years of the *International Journal of Aviation Psychology*. *International Journal of Aviation Psychology*, 10(1), 1–11.

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors*, 43, 641–674.

Tsang, P. S., & Vidulich, M. A. (2002). *Principles and practices of aviation psychology.* Mahwah, NJ: Lawrence Erlbaum.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Wickens, C. D., & Andre, A. D. (1999). Psychology applied to aviation. In A. M. Stec, & D. A. Bernstein (Eds.), *Psychology: Fields of application* (pp. 184–198). Boston: Houghton Mifflin.

Wilson, G. F. (Ed.). (2002). Flight psychophysiology [special issue]. *International Journal of Aviation Psychology*, 12(1).

Zsambok, C. E., & Klein, G. (Eds.). (1997). *Naturalistic decision making.* Mahwah, NJ: Lawrence Erlbaum.

Aviation

Evan A. Byrne

National Transportation Safety Board, Washington, D.C., USA

GLOSSARY

aviation psychology Scientific discipline focused on understanding and improving human performance in aviation.

crew resource management A collection of procedures, techniques, and communications strategies that fosters the use of all available resources (including people and available technology) to provide optimal levels of safety and mitigate the consequences of error.

cultural factors Procedural, communications, equipment, and design-related issues that are unique to an organization or a society.

electroencephalogram (EEG) Measurement of aggregate brain electrical activity recorded at the scalp.

FAA Abbreviation for Federal Aviation Administration.

NASA Abbreviation for National Aeronautics and Space Administration.

National Airspace System The complex system of people, equipment, and procedures that guides how aircraft are operated in the United States.

NTSB Abbreviation for National Transportation Safety Board.

situation awareness How people keep track of critical features in their environment, such as their position in space and time, and the current and future status of systems.

Aviation psychology is a simple name for a deceptively broad multidisciplinary applied domain. Its goal is plain: understanding and improving human performance in aviation. As a formal discipline, it draws on other established areas of psychology and engineering. For example, within the discipline of psychology, aspects of applied, experimental, social, clinical, industrial, organizational, and engineering specializations have been incorporated into aviation psychology.

1. HISTORY

The origins of aviation psychology are nearly as old as powered flight. Whether it played a role in the selection of the Wright *Flyer*'s first pilot on the sands of Kitty Hawk, North Carolina, a century ago is debatable. Lindbergh's self-account of his 33-hour transatlantic crossing in the *Spirit of St. Louis* in 1927 is among the earliest accounts of the effects of sustained wakefulness on pilot performance. However, aviation psychology is generally described as having its formal beginning coincident with the increased military use of aviation in the two world wars. During those times, there was a fundamental need to rapidly select and adequately train the pilots required to fly the fighters, bombers, and transport airplanes for the war efforts. As aviation became increasingly important to worldwide commerce, and with the advent of the space program, aviation psychology as a discipline expanded during the postwar years. This expansion paralleled technological

advances that increased the complexity of some facets of aviation and the continuing need to understand human performance capabilities and limitations in these domains. Several detailed recent reviews of the domain exist that expand on the history and basic applications in this area.

2. SCOPE AND EXTENT OF THE DOMAIN

Enhancing pilot performance is a major focus of aviation psychology; however, its reach in aviation extends far beyond pilots. For example, early efforts of aviation psychologists included enhancing the training of aircraft gunners and selection of aircraft mechanics. Aviation psychology focuses on understanding the operator's basic limitations and needs. Basic needs for optimal performance typically include adequate selection and training, sufficient real-time information, appropriate human–machine interfaces and equipment, and compensated environmental conditions. The scope of aviation psychology has generally kept pace with the rapid expansion of aviation itself since the advent of powered flight.

The National Airspace System is described on the Federal Aviation Administration's (FAA) Web site as "a complex collection of systems, procedures, facilities, aircraft, and ... people [that] represents the overall environment for the safe operation of aircraft." With its focus on the people and equipment involved in this complex system, aviation psychology is broad in scope but singular in purpose: to understand and promote optimal human performance. Aviation psychologists have contributed to our knowledge of human performance in both civilian and military applications. They have examined and sought ways in which to improve the performance of pilots, flight attendants, mechanics, air traffic controllers, and ground support personnel. It is clear that aviation psychologists now also have an important contribution to make in the study of human performance factors in civil aviation security, for example, in selecting and training security screening personnel and in identifying methods to augment and support human search capabilities during repetitive inspection tasks.

A peer-reviewed journal, the *International Journal of Aviation Psychology,* serves as a central focus for the dissemination of research in this area, and a survey of its contents can quickly provide a comprehensive overview of this area. Other professional journals, such as *Human Factors* (published by the Human Factors and Ergonomics Society) and *Aviation, Space, and Environmental Medicine* (published by the Aerospace Medical Association), also contain work by aviation psychologists. In the latter case, there are some who may be more appropriately titled "aerospace psychologists" given that their focus involves understanding human performance in space-related applications such as low-*g* environments and crew coordination issues during long-duration space flights.

3. SOURCES OF BASIC PRINCIPLES AND THEORETICAL APPROACHES

Aviation psychology draws on basic principles, findings, and tenets from other specialized domains of psychology, sociology, physiology, and human factors. For example, knowledge of sensory processes, memory limitations, psychomotor skills, and cognitive functioning translates directly to the aviation environment and can be used to optimize the design of instruments, procedures, and warning devices. Other areas of experimental and cognitive psychology that provide insight into attention and decision making also contribute to aviation psychology in areas such as evaluating instrument scan behavior, understanding visual detection capabilities for pilots (or inspectors detecting defects in manufacturing/maintenance inspection processes), and aeronautical decision-making skills. Finally, knowledge of basic human performance capabilities under exposure to noise, vibration, motion, and altitude facilitates aviation psychologists' understanding of human capabilities and limitations.

Methods derived from industrial, organizational, and educational psychology can aid in the selection and training of aviation personnel. Components from cognitive psychology, social psychology, and sociology can contribute to the study of crew coordination and cultural factors in aviation. For example, cultural factors in aviation are not just associated with design stereotypes for the movement of controls and switches. They also include crew coordination issues that can vary across cultures, include variations in the acceptance of automated technology, and can even affect the coordination of operating practices when airlines merge.

Physiological psychology has contributed to the measurement of workload and assessment of fatigue in aviation environments. Although few operational

examples of the former exist, the use of physiological measures to infer the state of operators has taken place in basic and applied aviation research environments. These include traditional measures such as heart rate and electroencephalogram (EEG). In addition, eye movements have been used to study operator scan patterns for complex displays or while performing aviation-related tasks.

There are few, if any, theoretical approaches unique to aviation psychology. Instead, basic theories generated elsewhere are applicable to basic and applied questions in aviation. Some theoretical or worldview approaches are especially salient as applied to the aviation environment. For example, information processing theories that outline and predict how operators process information from the environment and manage conditions of increasing workload are useful in evaluating aviation work environments to ensure that workload and capabilities are not exceeded. Also, during recent years, significant effort has been focused on understanding and establishing theoretical and methodological approaches to the study of situation awareness, that is, how operators learn about and keep track of critical features in their environment, such as their position in space and time, and the current and future status of systems. Theoretical models for the acquisition and maintenance of situation awareness by personnel involved in aviation tasks have been developed, although acceptance of their utility is not universal.

Aviation psychology presents a unique application and research environment. However, with few exceptions (e.g., operations in extreme environments, operations under conditions of g loading, specific examination of flight- or domain-specific tasks such as the visual processes involved in landing), knowledge, methods, and concepts derived elsewhere in psychology and engineering can readily translate to applications and research questions in this area. For additional information on the basic tenets, principles, and theoretical approaches used by aviation psychologists, there exist recent compendiums of work outlining these issues from both historical and contemporary perspectives.

4. APPLICATIONS AND ONGOING ISSUES

Aviation psychology is an incremental and evolving discipline that is generally responsive to the dynamic challenges it faces as aviation itself continues to evolve.

Following are some of the areas of inquiry in which aviation psychologists are engaged.

4.1. Error Management

Studies and training methods for crew coordination have evolved from attending static personality and "charm school" sessions to delivering strategies for the prompt identification and timely recovery from human errors that are inevitable. Work continues on determining optimal methods for enhancing crew coordination and error management, on developing instructional methods for distributing these error management concepts through technical training, and on reinforcing during routine operations through robust procedures.

4.2. Decision Making and High-Stress Situations

Research continues on understanding how personnel in the aviation environment arrive at decisions under conditions of stress, often having limited information. Some of these efforts are targeted to developing methods to enhance pilot decision making and judgment. Other efforts examine factors that affect human performance in stressful situations such as passengers evacuating an airplane under an emergency situation.

4.3. Fatigue Countermeasures

Substantial work in the area of fatigue has been performed by aviation psychologists at the National Aeronautical and Space Administration (NASA) Ames Research Center, and work continues to better understand the effects of fatigue on aviation operations. Research conducted by the FAA has looked at shiftwork and fatigue in air traffic control operations. Aviation psychologists working for the military have examined scheduling practices and pharmacological countermeasures to mitigate the effects of fatigue in military aviation.

4.4. Automation and Human–Machine Interfaces

The study of human interaction with automation remains a dominant focus for many aviation psychologists. The use of automated aids is increasing in both air carrier and general aviation, and this trend is expected to continue with the refinement of more complex aircraft

system status displays, horizontal and vertical navigation displays, airport situation awareness displays, highway-in-the-sky flight path aids, and the like. Automated aids are also increasingly being applied to other aviation operations such as air traffic control and maintenance.

4.5. New Training Technologies

Simulators have found widespread use in the training of air carrier flight crews. Methods for enhancing the training environment using these tools continue to be evaluated. The use of computer-based training and part task simulators is also spreading to other areas of aviation, including general aviation and maintenance technicians. Ensuring that the transfer of training effects is positive and predictable will continue to be a focus for aviation psychologists.

4.6. Accident Reduction

Methods to study human error in accidents continue to be examined to identify broad causal factors common to multiple accidents for which targeted intervention strategies could be applied. Strategies for this work include the use of surveys, the use of simulator studies, and review of accident data.

4.7. Maintenance and Other Aviation Applications

Some areas that have received relatively less attention in the past, as compared with flight deck and pilot issues, are receiving increasing attention by aviation psychologists. These include maintenance and ground support personnel, and recent accidents have shown that human error in these areas can also yield catastrophic outcomes. Strategies for increasing performance on the flight deck, such as crew resource management, are being translated to optimize performance by operators in these areas. Similarly, additional focus is being placed on the information needs of maintenance personnel to complement work done previously to examine environmental and ergonomics needs.

5. EDUCATION AND TRAINING

Because of the multidisciplinary characteristics of aviation psychology, there is no single training or education path to become an aviation psychologist (although for decades some institutions, such as the University of Illinois and Ohio State University, have produced professionals who made significant contributions in this area). Like the applied domain of human factors, successful performance in aviation psychology requires a professional who has developed a strong grounding in basic psychological principles, experimental methodology, and basic familiarity with engineering concepts. Creativity and the ability to adapt are essential traits that facilitate interaction with experts from the many facets of aviation, and formulation of strategies to cope with the short-cycle time periods (from question to required answer) characteristic of many issues aviation psychologists may be called on to study.

Equally important to becoming an effective aviation psychologist is a strong interest in, some observational experience in, or (optimally) direct experience in an operational facet of aviation (e.g., holding a pilot certificate). This experience will also help to facilitate communication with subject matter experts and provide the aviation psychologist with practical knowledge that will augment his or her professional training in a chosen area of psychology or in an allied discipline. Aviation has many regulatory and procedural boundaries that have been established by regulators, manufacturers, companies, and labor groups. These areas, whether they involve hours of service regulations for pilots or the maintenance and inspection routine, should be mastered to the extent possible for an aviation psychologist to contribute optimally.
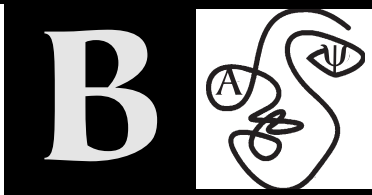
6. OCCUPATIONS AND EMPLOYMENT

Aviation psychologists can be found employed by organizations in and on the periphery of the aerospace industry. For example, they are employed by agencies such as NASA, the FAA, and the National Transportation Safety Board (NTSB) in the federal government and are involved in basic research, certification, design, oversight, and safety investigations. In addition, most branches of the U.S. Department of Defense (DoD) have similar areas in which aviation psychologists are employed to facilitate military aviation. In the private sector, aviation psychologists are employed by manufacturers, airlines, and private research institutions. Academia is another area where aviation psychologists can be found working on basic and applied research questions, often funded by DoD, NASA, and/or FAA grants and projects.

See Also the Following Articles

Accidents in Transportation ■ Engineering Psychology ■ Transportation Systems, Overview

Further Reading

Billings, C. (1997). *Aviation automation: The search for a human-centered approach.* Mahwah, NJ: Lawrence Erlbaum.

Caldwell, J. A., & Gilreath, S. R. (2002). A survey of aircrew fatigue in a sample of U.S. Army aviation personnel. *Aviation, Space, and Environmental Medicine*, 73, 472–480.

Dismukes, R. K. (1994). Aviation human factors research in U.S. universities: Potential contributions to national needs. *International Journal of Aviation Psychology*, 4, 315–331.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32–64.

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors*, 27, 149–157.

Garland, D. J., Wise, J. A., & Hopkins, V. D. (Eds.). (1999). *Handbook of aviation human factors.* Mahwah, NJ: Lawrence Erlbaum.

Guilford, J. P. (1948). Some lessons from aviation psychology. *American Psychologist*, 3, 3–11.

Helmreich, R. L., & Merritt, A. C. (1998). *Culture at work in aviation and medicine: National, organizational, and professional influences.* Aldershot, UK: Ashgate.

Koonce, J. M. (1984). A brief history of aviation psychology. *Human Factors*, 26, 499–508.

O'Hare, D., & Chalmers, D. (1999). The incidence of incidents: A nationwide study of flight experience and exposure to accidents and incidents. *International Journal of Aviation Psychology*, 9(1), 1–18.

O'Hare, D., & Lawrence, B. (2000). The shape of aviation psychology: A review of articles published in the first 5 years of the *International Journal of Aviation Psychology*. *International Journal of Aviation Psychology*, 10(1), 1–11.

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors*, 43, 641–674.

Tsang, P. S., & Vidulich, M. A. (2002). *Principles and practices of aviation psychology.* Mahwah, NJ: Lawrence Erlbaum.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Wickens, C. D., & Andre, A. D. (1999). Psychology applied to aviation. In A. M. Stec, & D. A. Bernstein (Eds.), *Psychology: Fields of application* (pp. 184–198). Boston: Houghton Mifflin.

Wilson, G. F. (Ed.). (2002). Flight psychophysiology [special issue]. *International Journal of Aviation Psychology*, 12(1).

Zsambok, C. E., & Klein, G. (Eds.). (1997). *Naturalistic decision making.* Mahwah, NJ: Lawrence Erlbaum.

Behavioral Assessment in Schools

Steven W. Evans, Amy L. Williams, and Brandon Schultz

James Madison University, Harrisonburg, Virginia, USA

Mark D. Weist

University of Maryland at Baltimore, Baltimore, Maryland, USA

GLOSSARY

analogue assessment An observational measure of the child's behavior(s) of interest elicited through the use of simulated conditions to predict behavior in real-life settings.

direct observation Observation and recording of problem behaviors in the context in which they occur.

effect size A calculation of preintervention and postintervention scores as well as the standard deviation or standard error of the instrument to produce an index of the amount of change assessed for the individual or group.

error variance The contribution of various sources of error (e.g., biases, drift) to a set of data.

functional behavioral assessment (FBA) Assessment of environmental conditions that are correlated with the occurrence and nonoccurrence of behaviors for the purpose of identifying those things that serve to initiate and maintain behavior.

generalization The degree to which learned skills or changes in behavior patterns are displayed in untrained situations.

informed consent The legal and ethical requirement to inform children and their legal guardians regarding the purpose of assessment procedures and to receive appropriate agreement to these procedures prior to their initiation.

interviews Direct questioning of the child himself or herself, or of the child's parents, teachers, and/or other caregivers, for the purpose of assessing strengths, weaknesses, and patterns of behavior; interviews can be highly structured (following a strict line of questioning), unstructured (based entirely on the interviewer's clinical direction), or semistructured (some combination thereof).

mediators/moderators Characteristics of the individuals receiving interventions and characteristics of the intervention process that affect treatment outcomes.

normative sample A group of people identified as being representative of a population who are assessed, with their data being used to compare with results obtained from subsequent administrations of the assessment instrument.

rating scales Assessment instruments that provide a standardized format to gather information about an individual's behavioral characteristics; these are usually paper-and-pencil forms that may be completed outside of an assessment setting.

response bias Trends in the way in which informants respond to assessments that result in error variance.

traditional assessment Behavioral assessment strategies that focus on assessing the presence or absence of diagnostic categories.

The assessment of social and emotional functioning in the schools is needed for purposes of classification, intervention development, and measurement of

outcomes. Procedures for completing these evaluations, as well as the relative strengths and weaknesses of the techniques, are described in this article. The descriptions are organized by purpose, beginning with techniques for problem identification and followed by the application of these techniques to measuring outcomes. Legal, ethical, and cultural issues as they relate to these procedures are described at the end.

1. RATIONALE FOR SCHOOL-BASED ASSESSMENT

There are two primary contexts for behavioral assessment in schools: special education and expanded school mental health (ESMH). When considering behavioral assessment in the context of special education, it is important to keep in mind that these services are guided in part by legislative acts such as the Individuals with Disabilities Education Act (IDEA) and the amendments to the IDEA in 1997 (IDEA '97). These legislative acts provide funds for public schools to fulfill the mandate of a "free and appropriate public education" for all children and adolescents with disabilities, including those with emotional and behavioral disabilities. In this context, behavioral assessment of students is critical for two reasons: (a) to assist in understanding how behavioral and emotional issues interfere with learning and (b) to develop appropriate intervention strategies to enhance learning success. Within IDEA '97, there is specific reference to the role of "functional behavioral assessment" (FBA) for the analysis and monitoring of student behavior. Although the law stops short of defining this requirement, FBA is widely assumed to include interviews, structured observations, and assessment of the function or purpose of student behavior. Based on the language of IDEA '97, FBA is intended to go beyond the mere question of eligibility for special education services or simple "gatekeeping." Rather, its intended purpose is to address those issues surrounding the creation and monitoring of appropriate interventions. To this end, IDEA '97 requires school professionals to assess behaviors that impede a student's ability to learn or that impede the ability for his or her peers to learn, although the language of the law does not specifically define those behaviors. Legal precedents based on case law suggest that physical aggression, noncompliance, verbal abuse, and disruptive classroom behaviors can rise to this level. For those students with disabilities who exhibit these types of behaviors, a decision must

be made as to the severity of the behaviors and whether or not they warrant specific interventions. If so, school professionals are legally bound to design a behavioral intervention plan (BIP) to address these needs. IDEA is being reauthorized in the U.S. Congress, and many hope that with this reauthorization there will be increased funding and an emphasis on prevention and support for youth before formal referral into special education. These changes would increase the need for assessment earlier in the process than is currently required.

Whereas IDEA '97 is applicable only to students in public schools who meet diagnostic criteria for specific areas of disability, Section 504 of the Rehabilitation Act of 1973 protects individuals from discrimination in any institution receiving federal funding, based on a much wider range of disabling conditions. Section 504 is not program legislation but rather civil rights legislation. In addition to its application to schools as a mechanism to receive special services, this act also requires that school districts ensure that all of their procedures and techniques are nondiscriminatory. When conducting behavioral assessment, professionals must be cognizant of potential cultural biases that may occur within the process. For example, behavioral assessment instruments that are standardized and norm referenced with samples that do not reflect the cultural differences of minority students may classify their behavior inaccurately. In such instances, it is imperative for school professionals to interpret results within the appropriate framework of the students' cultures and backgrounds. In other words, results from instruments that are based on norm samples that do not accurately reflect students' cultural backgrounds must be qualified. This is particularly true if these instruments have the potential for overidentifying minority students for disabilities based solely on cultural differences. Conversely, it is also important to not deny minority students special education services as a result of attempting to be culturally competent or politically correct.

The second primary context for behavioral assessment in schools is ESMH programs. ESMH programs involve collaboration between schools and community agencies to move toward a full continuum of mental health promotion, including early intervention and treatment for youth in special and regular education. These programs are developing progressively in the United States, as well as in other countries, as a result of their documented ability to reach underserved youth and preliminary evidence that they lead to outcomes valued by schools and communities (e.g., improved student attendance and behavior, decreased inappropriate referrals to

special education). Behavioral assessment is of foundational importance in the ESMH context and for special education and Section 504 plans to evaluate the nature of the presenting behavioral and emotional problems, provide guidance in the development of effective interventions, and measure the outcomes of those interventions.

2. EVALUATING BEHAVIORAL AND EMOTIONAL PROBLEMS

2.1. Purpose

The evaluation of behavioral and emotional problems in the classroom is warranted by the difficulties experienced by children who contend with these problems on a day-to-day basis. A number of behavioral and emotional problems that can prevent a child from functioning at a level commensurate with his or her same-age or same-grade peers are commonly evidenced in the school setting. Some of the more common behavioral and emotional problems that cause such impairments include anxiety, attention deficit/hyperactivity disorder (ADHD), anger/aggression, attachment disorders, conduct problems, depression, delinquency, posttraumatic stress disorder (PTSD), and bullying. The school environment is unique in that it places demands on children that are not otherwise required of them. For example, in the school setting, children are expected to sit quietly and attend to classroom activities. Furthermore, schools require students to function within demanding social contexts that youth may avoid or not be exposed to outside of school. The difficulties these expectations pose for children with behavioral and emotional problems are substantial, and assessment of these problems in the setting where they often occur has important implications for the development of effective interventions. It logically follows that assessment should occur, at least in part, in the school setting.

Evaluation of behavioral and emotional problems in schools is frequently done to determine whether a child or an adolescent is eligible for, or would benefit from, special services. Under these circumstances, evaluation is done to answer questions pertaining to classification and to guide the development of effective interventions. Although determining eligibility for special services is important, it is the development of interventions that is vital for changing problem behaviors. Effective interventions are those that target the presenting behaviors and not the label. Therefore, school professionals should place an emphasis on the development and implementation of interventions that target the specific needs of the students who are referred for evaluation rather than their diagnoses or labels.

2.2. Traditional Assessment

There are a number of assessment methods used to evaluate behavioral and emotional problems in the schools. Traditional assessment involves the examination of scores from norm-referenced measures to determine a particular individual's difference from or similarity to a normative sample. In other words, traditional assessment asks the question, "How does this student's behavior compare with that of similar children?" Standard tests of cognitive abilities and academic achievement, as well as many rating scales, are structured in this way. Using traditional assessment techniques, the inferred construct being measured (e.g., attention, depression) is thought of as a cause of a particular behavior in a specific setting. Although identifying the underlying syndrome or "label" related to a child's behavior is certainly important for diagnosis and classification, this information is not always the most useful for designing interventions. In contrast, measurement approaches such as behavioral observation, analogue assessment, interviews, and rating scales provide highly useful information about the functional relationship between the environment and particular behaviors. These techniques attempt to uncover the causes and consequences of behavior that often lead directly to effective intervention strategies.

2.3. Direct Observation

Direct observation is likely the most straightforward of all assessment techniques because it involves just what it implies: the direct observation and recording of problem behaviors in the context in which they occur. Three methods of direct observation are narrative recording (i.e., writing down what is happening in a narrative fashion, often following the antecedent–behavior–consequence [ABC] pattern or what occurs before, during, and after specific behaviors), event recording (i.e., tallying the instance of a target behavior), and time sampling (i.e., breaking down observations into intervals and recording the number of times a behavior occurs relative to the number of intervals observed to form a ratio). Direct observation differs from other methods of assessment in that little inference is required, making it one of the less biased methods for evaluating potential target behaviors.

There are some measurement issues related to observational methods that can negatively influence the data that are collected, including poorly trained observers, observer drift, poor operational definitions of target behaviors, poorly constructed data-recording systems, and reactivity (i.e., when the act of observation affects the student's behavior). Establishing interobserver agreement is one way in which to ensure that the data collected from direct observation are reliable and useful. A common procedure used to measure and enhance the quality of direct observation data is having another trained observer collect data simultaneously with the primary observer and then comparing data across observers to yield a measure of interobserver agreement. Direct observation is a labor intensive procedure that requires significant training time and effort to maintain reliability.

An additional problem with direct observation is the limited availability of the technique to assess infrequently occurring behavior. For example, children who are perceived as aggressive by teachers might exhibit only two or three episodes of these behaviors per week (or even fewer). The likelihood of observing these defiant or aggressive behaviors during discrete classroom observations is small. The technique is probably most practical for high-frequency behaviors such as on-task behavior and "talking out."

Finally, most observational data are difficult to put into context because there are no norms. In the school environment there is considerable variability in behavioral expectations across classroom activities and between teachers, resulting in various and shifting definitions of what is "normal." There are observational techniques that allow one to begin to account for this shifting context. For example, some time-sampling procedures require the observer to shift back and forth between observing the target child and same-sex classmates or a randomly selected peer. Pooling the data from the target and peer intervals provides the evaluator with a definition of what was normal during the context observed. However, no strategy will always work to overcome the problem of shifting contexts, and school professionals must interpret the results of their observations carefully.

2.4. Analogue Assessment

Analogue assessment provides an alternative to direct observation and is useful in situations where direct observation is not feasible. Analogue assessment is an observational measure of the child's behavior of interest elicited through the use of simulated conditions; observations are made in the simulated setting to predict behavior in real-life settings. Assessments of this type can be administered by paper and pencil, by audiotape or videotape, or through the use of role-plays or enactments. In a controlled environment, the evaluator creates a situation parallel to one that the child experiences in school and observes and records the behavior of interest. As with direct observation methods, this technique can be labor intensive; however, it does allow the evaluator to assess specific conditions that may occur only rarely. Some of these situations may be easy to recreate, for example, asking a child who is poor in math to complete some math and reading work. In this example, differences in the child's reactions to the various academic tasks may be observed, and these data may contribute to an understanding of behavior problems experienced during math class.

Among the problems associated with this method of assessment is the lack of standardization, issues related to reliability and validity, and problems with generalization. The most basic of these is the fact that analogue assessments are contrived situations, and regardless of how similar the situation is to the targeted environment, the child usually knows that it is contrived, and this knowledge may influence his or her behavior. Whenever analogue assessment techniques are used, all of these potential issues should be considered when interpreting the results.

2.5. Interviews

Interviews to assess behavioral and emotional problems in children can be conducted with the child himself or herself or with the child's parents, teachers, and/or other caregivers. A variety of different interview formats exist. An omnibus interview is designed for gathering a wide range of information, whereas a behavior-specific interview is narrower in scope and focuses on the assessment of specific problem areas. An interview that is problem solving in nature focuses on presenting concerns with the goal of developing an intervention plan. Interviews can be structured, unstructured, or semistructured.

Although it is often the case that interviews focus on reaching a diagnostic conclusion, interviews that focus on both problem assessment and the development of intervention plans are useful in school settings. Behavioral interviewing addresses the assessment of the current behaviors and environmental factors that contribute to the initiation and maintenance of problem

behaviors. These interviews are designed to detect specific problems that can be targeted for intervention and include phases of problem identification, problem analysis, treatment implementation, and treatment evaluation.

Interviews may be a valuable and efficient method for collecting information about a wide range of behavioral and emotional problems. The evaluator can efficiently collect data about behavior over a long period of time and in a variety of situations. In addition, the adults being interviewed may provide information about their reactions to this behavior or interventions related to the targeted behavior. Ideally, interviews are conducted by staff members who will continue to be involved in developing and implementing interventions with students in need. In these cases, the interview can be used to establish the initial relationship and communicate empathy and support to the child while collecting valuable information.

2.6. Rating Scales

Rating scale measures provide a standardized format to gather information about an individual's behavioral characteristics and can be administered to children to obtain self-report data or to parents, teachers, and/or other caregivers to obtain informant-report data. A few of the advantages of rating scales are their ability to provide behavioral information in a short amount of time, their moderate cost, and their technical precision and practical utility. Rating scales can be used in the comprehensive assessment of potential behavior problems and deficits in functioning for screening, diagnosis/classification, treatment selection/intervention planning, progress monitoring, and outcome evaluation.

As with most other forms of assessment, there are some problems associated with the use of rating scales. These problems typically fall under one of two categories: response bias and error variance. Response bias refers to trends in the way in which informants respond to rating scales. Some examples of response bias are halo effects (i.e., rating the student in a positive or negative manner based on characteristics unrelated to what is being measured), leniency or severity effects (i.e., the tendency to have an overly generous or overly critical response set), and central tendency effects (i.e., the tendency to rate down the middle and avoid end point ratings). Error variance refers to the differences among several ratings of the same individual. The different types of error variance include source variance (i.e., differences in ratings provided by various persons), setting variance (i.e., differences in ratings across

situations), temporal variance (i.e., inconsistency of behavior ratings over time), and instrument variance (i.e., differences in ratings using different instruments).

Although rating scales are valuable tools, it is frequently tempting to overrely on them due to their minimal costs. Although they are efficient, there are many questions they do not answer (e.g., the role of environmental contingencies). Although environmental concerns are critical to the development of interventions, this information is generally not provided by rating scales. Furthermore, rating scales are only as useful as the sources. For example, children are fairly poor raters of their own disruptive, off-task, and inappropriate behaviors; consequently, adults have been identified as a better source for this information. On the other hand, children appear to be better at reporting their internal affective states and, as a result, are generally better raters of their own depressive and anxious characteristics. However, although these are general trends that should be considered when interpreting rating scales, there is considerable individual variability.

2.7. Functional Behavioral Assessment

A primary goal of FBA is to identify environmental conditions that are correlated with the occurrence and nonoccurrence of behaviors. This type of assessment provides a systematic analytic method for identifying those things that serve to initiate and maintain behavior. FBA incorporates many methods of assessment to gain a complete understanding of an individual's behavior in a particular environmental context. Approaches to FBA can be indirect (i.e., gathering behavioral information from rating scales, interviews, etc.), direct (i.e., directly observing the behavior in the context of the natural environment in which it occurs), or experimental (i.e., generating hypotheses about the maintaining contingencies for the problem behavior and testing them with the individual serving as his or her own control). Completing these assessments is an iterative process of data collection, hypotheses generation, and collection of additional data. The hypotheses provide the specific questions that guide the subsequent data collection that either confirms or redirects the hypotheses.

The identification of the cause–effect relations in the interaction of an individual and his or her environment is integral to the development of interventions aimed at changing behavior. Because FBA focuses on the relationship between particular environmental circumstances

and behavior, it has the potential to contribute greatly to the development of effective interventions for problem behaviors. A successful FBA provides information about these environmental circumstances affecting the problem behavior and also provides a clear direction for treatment.

2.8. Current Issues

Because all of the assessment techniques discussed present their own unique advantages and limitations, the use of multiple assessment techniques with multiple informants is often considered the best method for evaluating behavioral and emotional problems. Although the use of multiple informants (e.g., parents, teachers, students) has the potential to provide a wealth of information, inconsistencies in data from one source and method to another is an issue that is often encountered in behavioral and emotional assessment. Such inconsistencies must be taken into consideration when assessing behavioral and emotional problems and especially when identifying behaviors to target for intervention. Frequently, careful evaluation relying on multiple informants using multiple methods will yield data that are contradictory and inconsistent, making interpretation difficult. Nevertheless, this comprehensive strategy is recommended to reduce the influence of specific sources of error and to produce the opportunity to recognize and interpret convergent themes used to generate explanations and interventions.

3. EVALUATING OUTCOMES

3.1. Purpose

As noted in the previous section, children and adolescents manifest many of their problems related to behavioral and emotional problems at school. This not only warrants school-based evaluation of these problems, it frequently results in school-based provision of services. These services take many forms, ranging from universal interventions, to disorder-specific treatment, to school-based day treatment programs. Professionals providing these services are expected to assess the progress of children and adolescents participating in them so as to modify interventions and know when interventions are no longer needed. As a result, the questions that guide outcome evaluations pertain to understanding an individual's progress in treatment with implications for modifying or discontinuing services. These assessment

procedures frequently include observations, rating scales, and performance measures.

3.2. Observations

Observations have inherent advantages and disadvantages that are described in the previous section. The use of observation measures to measure outcomes is warranted with high-frequency behavior or with behavior that has a high degree of situational specificity. Observed rates of on-task behavior is a common outcome measure of a high-frequency behavior. Defiant and disruptive behavior is usually difficult to evaluate with observational techniques because its frequency is low even for students who teachers perceive as defiant or disruptive. This is especially true in secondary schools, where episodes of defiant behavior can be quite intimidating and need not occur very often for a teacher to report a problem. Low-frequency behavior with a high degree of situational specificity is conducive to observational measures because the observer can efficiently target the specific situation that elicits the problematic behavior. For example, a child referred for problematic behavior in the cafeteria may warrant observational assessment because the observation period is limited and the likelihood of observing the problematic behavior within this situation may be quite high. Observational measures have been found to be sensitive to intervention effects when administered frequently and reliably.

3.3. Rating Scales

The use of rating scales can be a very efficient method for collecting outcome data. Compromises to validity can occur when raters are invested in the intervention process, resulting in a bias to perceive improvement. Collecting data from some raters who are blind to treatment or not invested in its success can be a valuable assessment strategy; however, sometimes these people are not available or are not able to provide useful assessment data. Collecting rating scale data from multiple sources and then interpreting convergent findings has been recommended so as to minimize individual biases. This technique can be helpful, but it also presents evaluators with unique challenges. It is quite common to receive rating scale data with tremendous variability and little convergence around any conclusion. Although this may indicate poor reliability, it may also indicate that the child's behavior changes as a function of classroom, activity, or teacher. In these situations, it is sometimes helpful to evaluate improvement in relation to

specific situations and settings. In other words, a child may be making considerable improvement in Mr. Smith's math class but demonstrating no improvement in Mrs. Martin's English class.

One of the problems with the use of rating scales as measures of outcomes pertains to the instruments possessing the necessary sensitivity to detect meaningful change. Changes resulting from effective interventions might not appear in changes in scores on rating scales because the targeted behavior change may have been more specific than can be indicated on most rating scales. For example, interventions may effectively reduce a set of behavior that annoys peers in math class; however, this may produce very little change in the child's overall social functioning or acceptance by other children and adults. As a result, the use of most rating scales as an outcome measure in this situation is probably inappropriate.

To measure changes in behavior, it is frequently useful to have repeated measurement of target behavior so as to track change. One of the practical problems with administering rating scales is the timely completion and return of assessments by parents and teachers. It is usually necessary to closely monitor their return, check for completion (e.g., whether the rater completed the back of the form), and follow-up with raters whose forms are incomplete or late.

3.4. Performance Measures

Performance measures assess the ability of a child or an adolescent to exhibit a skill. They are frequently used in academic achievement testing, but they also have applications related to behavioral and emotional problems. Pre- and post-assessments can be administered to assess whether a child had a skill (e.g., social skill, problem-solving skill) prior to the intervention and possesses the skill after the intervention. These assessments tend to be straightforward and based on an appraisal of the child's ability to demonstrate through role-play, written responses, and/or verbal description a set of skills that have been taught directly. These techniques are useful for determining whether a child or an adolescent is ready to begin applying the skills in actual situations because generalization is not likely to occur if the child cannot produce the behavior on cue. A common misapplication of these measures in evaluating outcomes is an overreliance on them as the end point of the need for interventions. Many socially impaired youth have been successfully taught numerous social skills that have resulted in no change in

behavior in other settings. Generalization, or the degree to which learned skills are displayed in novel or untrained situations, is usually the most challenging part of any behavioral intervention, and measures assessing behavior change in the target settings is a necessary outcome measure.

3.5. Current Issues

Successfully evaluating outcomes requires giving attention to many issues, including generalization, mediators and moderators, and size of effects. Attention to generalization requires evaluators to focus key outcome measures on the settings in which the problems that led to the initial evaluation occurred. Although measurement of the process of interventions can help to inform treatment decisions, it is possible to understand the ultimate success of interventions only when assessments target the problems in the context in which they have occurred. Moderators and mediators refer to characteristics of the individuals receiving interventions and characteristics of the intervention process that affect treatment outcomes. For example, interventions provided by different teachers in different classrooms may result in mixed results, not as a result of the intervention but rather as a result of whether or not each teacher implemented the techniques correctly. In this example, it could be said that the implementation of the technique "mediated" the outcomes.

Although data indicating improvements are encouraging, questions will remain as to how much improvement is necessary before it is concluded that meaningful changes have taken place. This issue is at the heart of indexes of effect size and clinical significance. These calculations rely on data from the same assessment techniques as described previously but include analyses to assist in the interpretations of the size of the effects. There are various approaches to this topic, including measures of whether change moved a student from outside of the normal range to within the normal range as well as indexes of the size of change in relation to the standard error of measurement and standard deviation. Unlike many other analyses, these techniques can be applied to data for an individual child or a group of children.

4. LEGAL AND ETHICAL ISSUES

For states to receive federal funding, there must be strict adherence to the principles of the applicable

program legislation and civil rights legislation passed at the federal level. State laws are often more specific than federal legislation and act to flesh out federal mandates, such as IDEA '97, by defining the means that school professionals use to meet the requirements outlined in these laws. With states operating independently in this fashion, there are often differences among states in terms of how the federal laws are interpreted. One state's procedures for conducting an FBA or a BIP may differ from those of another state. Furthermore, requirements set forth in IDEA that are not addressed by the state may be interpreted and defined at the school district level, sometimes resulting in dramatic differences in assessment and classification procedures between adjacent neighborhoods. Therefore, the onus is placed on schools to train their staff to be able to conduct behavioral assessment in accordance with the pertinent laws and local procedures.

In terms of behavioral assessment, perhaps the two most relevant areas of legal and ethical concern are informed consent and procedures for ensuring cultural competence in assessment. As with any form of non-academic assessment in the school setting, legal and ethical issues arise around consent for assessment. IDEA and the American Psychological Association's Ethical Principles of Psychologists specifically require that informed consent be obtained from the parents or guardians of a minor before conducting an initial evaluation. For consent to be "informed," the signor must (a) fully understand the purposes and means of the assessment, (b) provide consent voluntarily, and (c) be legally competent. It is generally assumed that children are incompetent to provide independent consent, although there are situations where exceptions are made. For example, students who are emancipated from their parents may be eligible to give consent on their own behalf, and many states allow independent consent for mental health services under 18 years of age. In general, however, parents or guardians are typically required to consent to any assessment services provided to children.

When obtaining consent, professionals must keep in mind that the school environment creates a unique situation for behavioral assessment because, in contrast to clinical settings, it is not automatically assumed that a child's participation will include such services. The clinical practice of "passive consent," or the assumption of consent based merely on participation, is not applicable in the school setting. In fact, some argue that the school setting is inappropriate for mental health services in general and

that the practice of assessing children in a school in domains outside of academia is inappropriate. Therefore, professionals in the school setting have the legal and ethical duty to ensure that parents or guardians understand the purpose and means of any assessment that is conducted. It is strongly recommended that school professionals obtain "active consent," which includes formal documentation of parental consent and meets all of the requirements of consent described previously. Ultimately, this might mean that students will not receive the services they need due to parental disapproval or apathy. Although this is unfortunate, the need for active consent presents such an ethical, and at times legal, concern that it is suggested that professionals in the schools always seek active consent despite the possible consequences.

Another area of legal and ethical concern is cultural competence. In simplest terms, there are two components of cultural competency. First, professionals must understand and respect the values, beliefs, customs, and traditions of the families they serve. This component of cultural competence involves interacting appropriately with people from diverse cultures, implementing a communication process whereby practitioners interact with people as equals (and not as experts), and demonstrating a genuine commitment to understanding students' strengths and weaknesses in the context of the culture, resources, and environment in which the behaviors occur. As mentioned previously, when a professional uses assessment instruments that are normed on nonrepresentational samples, there is a need to interpret the results with great care. Cultural competence in this respect may involve the ability to interpret the results within the correct framework. For example, a child of Hispanic background who appears to be significantly "introverted" or to lack "leadership" skills on some rating scales may in fact exhibit behaviors well within ranges that would be expected of his or her family-oriented cooperative culture. It is the culturally competent professional who can interpret these results in the correct context.

The second component to cultural competence is an awareness of how professionals' own cultural backgrounds affect their services to the students. In some cases, whole schools can adopt a white middle-class culture as a result of the dominant cultural makeup of the faculty and staff. In these instances, there may be a tendency within the school to place an emphasis on timeliness, parental involvement (to the exclusion of extended family members), and a lack of emotional display. Although this might not be self-evident

among the faculty and staff members of the school, these cultural norms can be at odds with families of diverse cultural backgrounds. For example, school personnel can mistakenly interpret students who arrive late for an assessment, or parents who arrive late to meetings, as uncaring or uninvolved. However, in a family's culture, there might not be the same emphasis on timeliness. Similarly, students who openly weep at unfortunate news or seem to be unmoved by positive feedback may be mistakenly considered to be unstable or unconcerned. Again, these reactions might not fit with the dominant culture of the school or professionals themselves, but they still fall well within the norm for the student's culture. As a result, professionals must continually assess their own cultural influences and be willing to recognize instances where their own cultural values do not match those of the students and their families.

See Also the Following Articles

Attention Deficit/Hyperactivity Disorders (ADHD) ■ Autism and Asperger's Syndrome ■ Cross-Cultural Psychology, Overview ■ Cultural Complexity ■ Cultural Psychology ■ Educational and Child Assessment ■ Emotional and Behavioral Problems, Students with ■ Ethics and Social Responsibility ■ Forensic Mental Health Assessment ■ Learning Disabilities ■ Psychological Assessment, Standards and Guidelines for ■ Psychology and the Law, Overview ■ School Discipline and Behavior Management ■ School Violence Prevention ■ Suicide Intervention in Schools

Further Reading

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/Adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232.

Dunlap, G., Kern-Dunlap, L., Clarke, S., & Robbins, F. R. (1991). Functional assessment, curricular revision, and severe behavior problems. *Journal of Applied Behavior Analysis*, *24*, 387–397.

Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, K. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review*, *30*, 193–210.

Ghuman, H. S., Weist, M. D., & Sarles, R. M. (Eds.). (2002). *Providing mental health services to youth where they are.* New York: Brunner–Routledge.

Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review*, *30*, 156–172.

Mash, E. J., & Terdal, L. G. (Eds.). (1997). *Assessment of childhood disorders* (3rd ed.). New York: Guilford.

Pellegrini, A. D. (2001). Practitioner review: The role of direct observation in the assessment of young children. *Journal of Child Psychology and Psychiatry*, *42*, 861–869.

Shapiro, E. S., & Kratochwill, T. R. (Eds.). (2000). *Behavioral assessment in schools: Theory, research, and clinical foundation* (2nd ed.). New York: Guilford.

Sterling-Turner, H. E., Robinson, S. L., & Wilczynski, S. M. (2001). Functional assessment of distracting and disruptive behaviors in the school setting. *School Psychology Review*, *30*, 211–226.

Weist, M., Evans, S. W., & Lever, N. (Eds.). (2003). *Handbook of school mental health: Advancing practice and research.* New York: Kluwer/Plenum.

Behavioral Medicine Issues in Late Life—Pain, Sleep, etc.

Bruce Rybarczyk

Rush Medical College, Chicago, Illinois, United States

GLOSSARY

depressed psychophysiologic functioning A complication from hospitalization (increased incontinence, confusion, agitation, and loss of appetite) often observed in the frail old.

excess disability A substantially lower than expected level of physical functioning for a given medical condition, attributable to behavioral factors (i.e., noncompliance, lack of activity, and depression).

frail elderly The old-old (age 85+ years) segment of the population with significant, multiple medical illnesses that limit functioning. These individuals typically do not live independently.

masked depression A presentation of depression in older adults characterized by the presence of more physical symptoms (i.e., reduced appetite, sleep disturbance, fatigue, and increased aches and pains) and fewer mood symptoms.

pseudodementia Reversible symptoms of dementia (i.e., memory, attention, and other cognitive deficits) that are caused by an underlying depression.

sleep hygiene An educational treatment for insomnia that provides instruction on how the sleep environment, caffeine, alcohol, nicotine, food, and exercise affect sleep.

sleep restriction A treatment for insomnia that involves restricting the time allowed for sleep, beginning with a very short period and gradually increasing. This method promotes a state of sleep deprivation in order to facilitate learning to fall asleep faster and stay asleep.

stimulus control A treatment for insomnia that requires the individual to leave the bed when he or she is not able to sleep and to not use the bed or bedroom for any activities during the day.

Behavioral medicine encompasses the study of how behavior impacts the onset, progression, and recovery from illness and the application of those behavioral factors to promoting health and preventing illness. Due to national demographic trends and the prevalence of chronic diseases in late life, older adults are emerging as a critical target group for behavioral medicine interventions. In the older adult age group, behavioral factors relating to comorbid depression, excess disability, pain, and sleep are of particular importance for behavioral medicine.

1. APPLYING BEHAVIORAL MEDICINE TO OLDER ADULTS

Due to the twin factors of increased life expectancy and reduced birth rates, most nations throughout the world are experiencing a dramatic shift toward an older adult population. It is projected that by 2030, 70 million Americans (20% of the U.S. population) will be 65 years or older, compared to approximately 35 million (14%) in 2000. This aging revolution has led to increased public and scientific interest in how behavioral and psychological factors influence the aging process. In addition, with the great strides made in the successful treatment of acute illness in the past century, people are living longer but at the same time are accruing more age-related chronic illnesses. In response, the focus of the U.S. health care system is shifting toward managing chronic illness and disability, with an emphasis on maximizing independence and quality of life in late adulthood. The goal is to help people live well rather than simply live longer.

By necessity, the field of behavioral medicine will play an increasingly important role in the health care system of the future. Behavioral medicine encompasses the study of how behavior impacts the onset, progression, and recovery from illness and the application of those behavioral factors to promoting health and preventing illness.

Older adults are in many ways the ideal age group to apply behavioral medicine principles and interventions. First, they are at much greater risk for chronic medical conditions, such as heart disease, lung disease, diabetes, cancer, and arthritis, and they often have multiple chronic conditions. Interventions aimed at preventing the onset or progression of such conditions are most cost-effective when they are designed for and delivered to this age group. In-home exercise interventions designed specifically for older adults, for example, have been shown to yield substantial benefits to patients with cardiovascular disease, partly due to the fact that older adults are more vulnerable to diminished cardiovascular capacity when they do not engage in systematic exercise.

Second, each chronic medical condition typically requires several different medications, multiplying the risk for side effects (e.g., memory impairment) and drug interactions. Accordingly, behavioral interventions that can manage symptoms without the use of medications are particularly important for this population.

Last, and perhaps most important, is the fact that physical and psychological conditions become more interdependent and reactive to each other with advancing age. Therefore, interventions aimed at psychological conditions, such as depression and anxiety, as well as the enhancement of psychosocial coping resources can have a significant impact on the physical health of older adults. Mind–body group interventions that teach a range of skills (e.g., relaxation training, problem solving, and cognitive approaches to managing stress) have been shown to be effective in enhancing coping and quality of life in older adults with chronic illness.

2. DEPRESSION AND HEALTH IN OLDER ADULTS

The greater interplay between psyche and soma (body) in older adults is perhaps most evident in research on depression and health. When older adults have a depression overlaid on a medical condition, they are at much higher risk for excess disability. Excess disability is defined as a substantially lower than expected level of physical functioning for a given medical condition, attributable to behavioral factors (i.e., noncompliance and lack of activity). After a stroke or hip fracture, for example, when depression is detected an aggressive psychological intervention is required to prevent poor physical rehabilitation outcomes.

Even mortality has been shown to be affected by depression in older adults. In the Cardiovascular Health Study, depressive symptoms were related to 6-year mortality in a large sample of adults older than age 65, after controlling for medical and demographic variables known to affect mortality. The hypothesized mechanism by which depression affected mortality was decreased motivation to initiate and sustain cardiac health behaviors.

Unfortunately, medical professionals often fail to detect the presence of depression in older adults. This failure in detection is partly accounted for by the atypical presentation of depression in many older adults. One such presentation is as masked depression, which is characterized by the presence of more physical symptoms (i.e., reduced appetite, sleep disturbance, fatigue, and increased aches and pains) and fewer mood symptoms. Furthermore, depressed older adults are far less likely to have a negative view of themselves (e.g., "I am a failure" and "I am worthless"), which is a hallmark of depression in other stages in life. Another presentation, called pseudodementia, is characterized by reversible symptoms of dementia

(i.e., memory, attention, and other cognitive deficits) that are caused by an underlying depression.

Fortunately, cognitive–behavioral and other psychotherapy treatment approaches have been found to be as effective as medication in the treatment of depression; moreover, older adults are as likely to benefit from these treatments as younger adults. Furthermore, there is no evidence to suggest that when chronic illness coexists with depression that psychotherapy treatment for depression is any less effective, although the situation is likely to contribute to the range of cognitive errors and distortions that need to be addressed in cognitive–behavioral therapy (Table I).

A positive development is a growing trend toward improved screening and treatment of depression in older medical patients as well as a greater presence of mental health professionals in primary care settings. This trend is particularly important given the fact that older Caucasian males are at the highest risk of any group for suicide and as many as 70% of suicide victims visit their primary care physicians prior to committing suicide. Tragically, these older male patients were unable to communicate their depression to their physicians and/or their physicians failed to ask the appropriate questions. Once depression is detected, psychotherapy that is provided on-site in a primary care office is often more acceptable to older adults. This detection treatment model highlights the importance of integrating mental health and medical treatment for older adults in particular.

TABLE I
Common Cognitive Errors and Distortions in Depressed Chronically Ill Older Adults

Irrational belief: disability = inevitable misery
  "Anyone would be depressed in my shoes."
All-or-none thinking
  "If I can't do everything that I used to do, in the ways I was accustomed to, I might as well not do anything . . . because it won't be satisfying."
Negative filtering
  "I'm nothing but a burden to my family."
Negative forecasting
  "If I continue to have all these medical problems, my friends won't want to be around me anymore."
Mind reading
  "The medical staff is getting tired of working with me."

3. BEHAVIORAL SLEEP MEDICINE AND AGING

Sleep has emerged as an important area in behavioral medicine. Treatment of insomnia, in particular, has been an important area of research and intervention because it is frequently caused or made worse by behavioral factors that are amenable to cognitive–behavioral intervention. In this case, again, the older adult population may be the ideal target group for these interventions.

There are numerous sleep disorders that increase in frequency with age, including sleep apnea, restless legs syndrome, and periodic limb movement disorder, but chronic insomnia is the most prevalent in late life. The risk for developing chronic insomnia doubles between the ages of 55 and 85 and is present in 20–30% of the population older than age 65. It usually manifests itself as a combination of different problems, including difficulty falling asleep, excessive time spent in bed awake, more frequent and longer awakenings after sleep onset, and fewer total hours of sleep. Furthermore, there is an increasing recognition that sleep disturbance at night leads to daytime fatigue and/or sleepiness, which is associated with decreased quality of life in older adults.

During the past 30 years, several different behavioral interventions have been developed and shown to be highly effective. In fact, it has been noted that behavioral interventions for insomnia are probably the most effective and durable interventions in the behavioral medicine repertoire, with typical success rates after only six to eight sessions of treatment of 60% or greater. In addition, most of these treatment effects have been measured and sustained over periods of as long as 2 years.

The most effective cognitive–behavioral interventions tested to date include all or most of the following components: sleep hygiene instruction, relaxation training, sleep restriction, stimulus control, and cognitive modification. Sleep hygiene involves basic education on how the sleep environment, caffeine, alcohol, nicotine, food, and exercise affect sleep. Several different types of relaxation training (e.g., progressive muscle relaxation, deep breathing, and visual imagery) have been provided under the premise that mental and physical arousal that accompanies stress and anxiety will inhibit and disrupt sleep. If patients can train themselves to relax before sleep or at night after an awakening, they are more capable of falling asleep and staying asleep.

3.1. Sleep Restriction

Sleep restriction is the most difficult to adhere to but is probably the most essential component of treatment of insomnia in older adults. Its aim is to counter the sleep fragmentation and daytime sleeping that is commonly observed in older adults by restricting the time allowed to sleep to a specific "sleep window." This sleep window is initially set for a very short period of time, determined by the average amount of time that the individual slept per night (based on a sleep diary that is kept throughout treatment). For many individuals, this period is 5 or 6 h long, so they are required to be awake for as many as 19 h each day prior to going to bed. During the first few weeks the individual experiences a high level of sleep deprivation, compounded by the fact that he or she is not yet capable of sleeping all the time during the sleep window. This prolonged sleep deprivation creates "sleep pressure," which makes it easier for patients to train themselves to sleep without interruption during the sleep window.

After each week of sleep restriction treatment, the patient's sleep efficiency (percentage of time slept during the sleep window) is calculated and 30 min is added to the sleep window if the patient achieves a sleep efficiency of 85% or greater. The eventual goal is to give the patient enough time in bed to meet his or her particular sleep needs (i.e., enough sleep to feel refreshed and not sleepy during the day) while maintaining a high level of sleep efficiency.

3.2. Stimulus Control

The stimulus control method is designed to increase the association between being in bed and being asleep. Predicated on operant learning theory, it involves instructing individuals to greatly limit their activities in the bed and bedroom. Basically, the only activities allowed in bed are sleep and sex. They are not permitted to do such activities in bed as watching television, reading, eating, worrying, listening to the radio, working, or simply resting. In addition, they are not allowed to be awake in bed for longer than 15–20 min at any time during the night, whether they are trying to fall asleep or trying to get back to sleep after awakening in the middle of the night or early morning. If they cannot fall asleep or back to sleep after 15–20 min, they are to get out of bed and engage in a nonstimulating activity until they become sleepy enough to try again. If they are unable to fall asleep again within 15–20 min, they are to repeat the same procedure. In addition, to facilitate the functioning of the body's circadian clock system, they are instructed to use an alarm clock and get up at the same time every morning.

The final component of treatment is cognitive modification. Similar to cognitive restructuring for anxiety or depression, this approach emphasizes changing unrealistic beliefs and irrational fears involving sleep or sleep loss. Research shows that older adults with insomnia endorse stronger beliefs about the negative consequences of insomnia, express more fear of losing control of their sleep, and express more helplessness about its unpredictability compared to older adults without insomnia.

Dysfunctional cognitions about sleep create a vicious cycle by increasing performance anxiety about sleep, which in turn leads to arousal and greater difficulty falling asleep. Common misconceptions about sleep include the following: blaming daytime impairments, such as fatigue and inability to concentrate, exclusively on sleep loss; the fear that insomnia will worsen a medical condition; the belief that not getting 7 or 8 h a sleep per night will threaten one's health; and the belief that sleeping late the next morning or napping is a good way to compensate for poor sleep at night. Finally, the belief that a bad night of sleep results in a "lost day" leads to cancellation or avoidance of activities, such as socializing and physical activities, that are important to maintaining health and promoting sleepiness at night.

Initially, the focus of research on behavioral treatment for insomnia was on younger adults and later it moved to older adults who had no other medical conditions. Until recently, individuals with chronic illness were excluded from these studies under the presumption that insomnia in these individuals was likely caused by a medical condition or a medication and therefore needed to be treated by directly managing the medical condition. However, it is now understood that insomnia in virtually all cases, regardless of coexisting medical conditions, is made worse by behavioral factors that are responsive to behavioral treatment. Indeed, two studies have demonstrated that older adults with multiple medical conditions are as responsive to treatment as individuals without medical conditions.

Insomnia in older adults occurs mostly in those with age-related chronic diseases, such as cardiac disease, arthritis, and lung disease. When insomnia coexists with chronic disease, these patients have significantly reduced quality of life, experience more pain, use more medical services, and are more likely to develop depression. Thus, providing behavioral treatment to ameliorate insomnia in these individuals has numerous benefits that go beyond improvement of sleep, leading to improved daytime mood, alertness, and physical functioning.

4. CHRONIC PAIN INTERVENTIONS FOR OLDER ADULTS

Another important application of behavioral medicine to the older adult population is the management of chronic pain. Pain is a constant factor in the lives of a great many older adults. Due to the greater prevalence of conditions that cause pain (e.g., arthritis and diabetes), older adults have more chronic pain and are more likely to have multiple sites of pain than younger adults. Pain in older adults has been strongly linked to excess disability, insomnia, depression, and increased health care utilization. Older adults with a combination of pain and depression have been shown to be particularly vulnerable to physical disability.

The two types of behavioral treatments for chronic pain that have been applied to older adults are cognitive–behavioral therapy (CBT) and operant conditioning. CBT methods focus on modifying beliefs, thoughts, attitudes, and skills that affect different aspects of the pain experience. They also employ relaxation training to reduce the sympathetic nervous system component of the pain experience. Operant conditioning focuses on modifying observable behaviors, such as activity level, medication usage, and pain signaling (i.e., grimacing, limping, and complaining). This method uses operant conditioning principles to reward activity, change environmental contingencies, and shift medication usage from an as-needed to a time-contingent approach. Both treatment methods use learning techniques such as consistent self-monitoring, graded practice, and homework. CBT and operant techniques are compatible and are often used together in clinical practice.

Contrary to traditional bias among professionals that older adults may not benefit from CBT interventions due to the introspection that is involved, studies demonstrate that older adults obtain equal benefits from CBT pain interventions compared to younger adults. In addition, they are no less likely to drop out of treatment. Older adults have also been shown to be as capable of learning relaxation procedures as younger adults. Nonetheless, research exploring age bias in inpatient and outpatient pain treatment programs has found that rehabilitation professionals remain less confident that older adults will benefit from their program. This bias is confirmed by the much smaller than expected percentage of older adults enrolled in these programs, which means they are being underserved by these invaluable treatments.

Older adults may also have some unique, age-related strengths and resources that help them cope better with pain. The majority of studies suggest that older adults have a higher pain threshold compared to younger adults. In addition, older adults may be more tolerant of pain due to the fact that they are more willing to accept certain levels of pain and discomfort as a part of growing older. This is consistent with studies showing that older adults report diminished emotional responses to pain, such as depression, anxiety, anger, or fear. Similarly, according to the life span theory of control, as individuals move into late adulthood there is a shift from a focus on primary control (control over the external world) toward an emphasis on secondary control (control of the internal mental and emotional life). This may equip older adults with a greater skill to cope with stressors that are irreversible or cannot be easily controlled, such as chronic pain.

5. BEHAVIORAL MEDICINE ISSUES IN THE FRAIL OLD

Among the old-old (age 85+ years) with significant medical illnesses, sometimes called the frail old, behavioral medicine has perhaps an even greater role to play due to the even stronger interplay between psyche and soma. When the frail old develop new medical problems that compound existing ones, they often experience the onset of new psychological symptoms, such as delirium, depression, or paranoia. Similarly, older adults residing in nursing homes have been shown to experience increased mortality and medical illness, not accounted for by medical factors, when they are required to relocate from one institutional setting to another. Even "timing of death" research has shown that, remarkably, the old-old are more likely to die of natural causes after, rather than before, a holiday or important family event. This suggests that the psyche exerts some degree of control over the timing of death.

When frail older individuals are hospitalized they often experience depressed psychophysiologic functioning in the form of increased incontinence, confusion, agitation, and loss of appetite. These complications lead to further medical interventions, such as catheterization for incontinence, the administration of sedating medications or use of restraints for agitation, or tube feeding in the case of failure to eat. These treatments, in turn, lead to iatrogenic effects (i.e., complications from treatment) such as bladder infections (from catheterization), blood clots (from being restrained to the bed), and pneumonia (from tube feeding).

6. SUMMARY

Due to the increased risk of chronic illness and the increased interaction between psyche and soma in late life, behavioral medicine interventions that target older adults are among the most cost-effective and essential. Interventions that reduce anxiety and depression or enhance coping in older adults with chronic diseases have much potential to yield crossover effects on medical variables, such as functional status, symptom severity, and cognitive functioning. These factors combined with the aging revolution in the United States will lead to an increased focus within behavioral medicine on providing services to older adults.

See Also the Following Articles

Aging, Cognition, and Medication Adherence ■ Depression in Late Life ■ Elder Caregiving ■ End of Life Issues

Further Reading

Rybarczyk, B., Gallagher-Thompson, D., Rodman, J., Zeiss, A., Gantz, F., & Yesavage, J. (1992). Applying cognitive–behavioral psychotherapy to the chronically ill elderly: Treatment issues and case illustration. *International Psychogeriatrics*, *4*(1), 127–140.

Schulz, R., Beach, S. R., Ives, D. G., Martire, L. M., Ariyo, A. A., & Kop, W. J. (2002). Association between depression and mortality in older adults: The Cardiovascular Health Study. *Archives of Internal Medicine*, *160*(12), 1761–1768.

Stepanski, E., Rybarczyk, B., Lopez, M., & Stevens, S. (2003). Assessment and treatment of sleep disorders in older adults: An overview for rehabilitation psychologists. *Rehabilitation Psychology*, *48*, 23–36.

Yonan, C. A., & Wegener, S. T. (2003). Assessment and management of pain in the older adult. *Rehabilitation Psychology*, *48*, 4–13.

Behavioral Observation in Schools

John M. Hintze

University of Massachusetts, Amherst, Massachusetts, USA

GLOSSARY

frequency or event recording Involves counting the number of times a behavior is observed during a specified time period.

interobserver agreement Technique in which the consistency of recording across two or more observers is compared to assess the reliability of measurement.

interval recording Divides the observational session into a number of equal intervals and simply records the presence or absence of specified behaviors during each interval.

narrative recording Refers to the observation and collection of information on student behavior in naturally occurring arrangements with little constraints placed on the observer for how and/or what to observe and record.

partial-interval recording An interval recording technique in which the interval is scored if the target behavior is present during any part of the interval.

systematic direct observation Refers to the observation of behavior that is explicitly defined under predetermined settings.

whole-interval recording An interval recording technique in which the interval is scored if the target behavior is present throughout the entire interval.

Behavioral observation refers to a method of assessment whereby human observers objectively record the ongoing behavior of a person or persons in specific environmental circumstances.

1. INTRODUCTION

Direct observation is one of the most widely used individual assessment procedures by school-based professionals. In a survey of more than 1000 school-based professionals, of the 26 different types of individual assessment instruments listed across seven different assessment categories (e.g., aptitude, social–emotional, and personality), behavioral observation methods ranked highest in terms of frequency of use. Overall, respondents indicated that on average they conduct approximately 15 behavioral observations during the course of a typical month.

2. GENERAL METHODS OF BEHAVIORAL OBSERVATION

When conducting behavioral observations, school-based professionals generally rely on one of two basic approaches to gathering information. One form, narrative recording, refers to the observation and collection of information on student behavior in naturally

occurring arrangements with little constraints placed on the observer for how and/or what to observe and record. With narrative recording, the observer generally describes the observed situation making notes of specific behaviors as they occur. The method(s) of observation, the type of information noted, and the manner in which the information is summarized are left to the discretion of the observer. The strength of such procedures lies in the flexibility that the observer has in choosing when and how to observe and the minimization of obtrusiveness or reactivity that may occur as a result of the presence of the observer. For interpretative purposes, narrative recording is often used to help develop hypotheses about the various behavioral and environmental factors that may be worthy of further observation and analysis. This flexibility, however, is also one of the main weaknesses of narrative recording. Because the observer has great autonomy in choosing how and what to observe, judgments about the worth of such reports are inextricably bound to the subjective judgments of the observer. As such, it would be highly unlikely for two independent observations collected at the same time to appear identical in the information provided. Therefore, narrative recording is generally best used as a precursor to more specific, objective accounts of behavior.

In contrast to narrative recording, systematic direct observation refers to the observation of behavior that is explicitly defined under predetermined settings. Although this approach is also concerned with observing behavior under naturally occurring environmental contexts, the aim is to define beforehand the behaviors of interest, choose specific recording strategies, and have observers record whenever behavior corresponding to the predefined operational definitions occurs. In particular, systematic direct observation is distinguished by five characteristics: (i) The goal of observation is to record specific behaviors, (ii) the behaviors being observed have been operationally defined a priori in a precise manner, (iii) observations are conducted using standardized procedures and are highly objective in nature, (iv) the times and settings for observation are carefully selected and specified, and (v) scoring and summarizing of data are standardized and do not vary from one observer to another. In comparison to narrative recording, a main objective of systematic direct observation is to have independent observers agree on what is observed and recorded, assuming that they have observed the same stream of behavior.

3. MEASURING AND RECORDING BEHAVIOR SYSTEMATICALLY

Various types of data can be collected during systematic direct observation. A workable definition of a target behavior is one that provides an accurate description of the behavior and clearly defines the boundaries of its existence and nonexistence. Therefore, constructs and reifications do not lend themselves well to direct forms of observation. For example, raising one's hand to be called on is an observable and measurable behavior. Behaving "off the wall" is not something that can be directly observed (although operational definitions of the behaviors that constitute "off the wall" could be developed). Once behavior is defined, the calibration of the operational definition is determined by the nature of the data—the frequency of its occurrence and the particular interests of the observer. In addition, practical considerations, such as the availability of observers, the amount of time a student is accessible, or any combination of these factors, all dictate the type of data collected.

4. CODING SCHEMES

When observing behavior systematically, observers generally use any one of a number or a combination of approaches to recording behavior. The more common approaches usually involve counting the number of times a behavior occurs during a specified time period and/or noting the presence or absence of behavior at specific time intervals of an observational session. Frequency or event recording involves counting the number of times a behavior is observed during a specified time period. When the time periods vary across multiple observational sessions, frequencies are converted to rate of behavior per unit time. For example, an observer may report that a child got out of his or her seat at an average rate of one time per minute during three separate observations conducted over the course of 3 days, even though the actual duration of each observation period varied. By using rate of behavior rather than frequency, comparisons can be made across observational sessions that differ with respect to time. Frequency recording is most useful for observing behaviors that have a discrete beginning and ending and are relatively consistent in the length of time that they take to occur. For example, "tantruming" may not lend itself well to frequency recording because the beginning and end of each tantrum might be difficult

to discern and the length of time of each tantrum might vary from a few minutes to hours. Moreover, frequency recording is usually better suited for behaviors that occur at lower rather than higher rates. With higher rates of behavior, accurately detecting each instance of behavior can be challenging.

In addition to frequency recording, another commonly used recording schedule is interval recording. In comparison to frequency recording, which notes each occurrence of behavior, interval recording divides the observational session into a number of equal intervals and simply records the presence or absence of specified behaviors during each interval. For example, a 20 min observation session could be broken down into 120 10 s intervals. Moreover, unlike frequency recording, in which each instance of behavior is noted, interval recording is only concerned with whether the targeted behavior occurs during the interval. As such, distinctions are not made regarding how many behaviors were observed during the interval. For example, an interval in which there were nine instances of out-of-seat behavior would be coded the same as an interval in which only one instance of out-of-seat behavior occurred.

In practice, interval-recording data are generally collected using one of three coding techniques. When using whole-interval recording, an interval is scored when the target behavior was present throughout the entire interval. Since the behavior must be present for the entire interval, whole-interval recording is well suited to behaviors that are continuous or intervals of short duration. One of the drawbacks of whole-interval recording is that it tends to underestimate the presence of the behavior. For example, if "off-task" behavior were the target behavior of interest and it was observed to have occurred for 8 s of a 10 s interval, the interval would not be scored for the presence of off-task since it did not occur for the entire 10 s. In contrast to whole-interval recording, partial-interval recording codes the occurrence of a behavior if it occurs during any part of the interval. Partial-interval recording is a good choice for behaviors that occur at a relatively low rate or for behaviors of somewhat inconsistent duration. However, because the interval is scored for any presence of the target behavior, partial-interval recording tends to overestimate the actual occurrence of behavior. Lastly, momentary time sampling codes the presence or absence of the target behavior at only one predefined instant of the interval (e.g., usually the instant in which the interval begins). As such, the target behavior is coded as occurring or not on the basis of a fractional observation of the total interval. Once the target behavior is coded, no other behaviors are noted for the balance of the interval. Although momentary time sampling appears counterintuitive because it is based on the smallest sample of behavior, it actually provides the least biased estimate of behavior, generally providing neither over- nor underestimates of actual behavior. In addition to frequency and interval recording, other less commonly used coding methods include duration (i.e., the actual amount of time that a behavior occurs), latency (i.e., the time it takes for a behavior to be initiated following a prompt or directive), and intensity (i.e., the amplitude of behavior).

5. OBSERVATIONAL INSTRUMENTS

In addition to observer-designed coding schemes, observational instruments have been developed to assess a specific range of behaviors in specific environmental circumstances. For example, an observer might choose to use an observational instrument designed specifically to quantify the percentage of time a student is academically engaged in or on-task, the frequency with which the student interacts with other students on the playground, or the number of times a teacher provides directives, opportunities to respond, or positive reinforcement. In comparison to observer-developed coding schemes, observational instruments can be somewhat more limited in their flexibility since they are usually developed with specific standardized operational definitions and recording schedules. On the other hand, since observational instruments are usually developed with a specific purpose in mind, they often provide a more detailed account of the student's behavior under the specific environmental context of interest.

6. SUMMARIZING BEHAVIORAL OBSERVATION DATA

Once behavioral observation data are coded they are usually summarized across the entire observational period to provide a description of behavior. As previously indicated, frequency data are reported as a function of rate (i.e., number of behaviors noted divided by the amount of time observed). By doing so, summary data collected across observational sessions of varying time can be compared. Interval data are expressed as a percentage of the intervals that the target behavior occurred compared to the total number of intervals observed. In summarizing

and reporting interval recording results, it is important to note that the behavior was coded for a certain percentage of intervals observed and not time (e.g., "Off-task behavior was noted for 60% of the intervals observed," not "Off-task behavior was noted 60% of the time"). This subtle distinction highlights that the coding scheme used was interval recording, and that intervals and not actual time was the unit of analysis. Duration recording is generally reported as the cumulative amount of time that a behavior was observed. However, it is not uncommon to record the duration of each instance of observed behavior separately and then sum the durations for a final cumulative total. In addition to providing total duration, the latter procedure can also provide the average duration per instance of behavior. Latency recording is summarized similarly to duration, with either cumulative latency or average latency per behavior generally noted. Lastly, intensity is usually summarized using some form of subjective ordinal scale of analysis (e.g., spoke in a voice loud enough to be heard by a person sitting next to him or her, spoke in a voice loud enough to be heard by a person in the room, spoke in a voice loud enough to be heard by a person outside the room, etc.) or by some mechanical recording device specifically designed to measure the dimension of interest (e.g., amplimeter).

7. RELIABILITY AND VALIDITY OF BEHAVIORAL OBSERVATION

The applicability of reliability and validity to behavioral observation has been equivocal, with some suggesting that classical psychometric concepts based on differences between persons are irrelevant to an assessment methodology that focuses on behavior and its variation within individuals, and others suggesting that the differences between traditional and behavioral assessment are primarily conceptual, not methodological, and as such reliability and validity considerations apply.

Traditionally, behavioral assessors and researchers have approached the issue of reliability using accuracy and interobserver agreement as substitutes. Specifically, accuracy refers to the extent to which observed values of behavior approximate the "true" state of the behavior as it actually occurs. When the true state of the behavior is known, comparing the observed values to the true values derives accuracy of measurement. In this case, the behavioral notion of accuracy is the same as reliability in the classic measurement case in which reliability is represented by the extent to which an observed

score represents the true score. As such, in order for a measurement system to be accurate it must be sensitive to the occurrence of true behavior, its repeated occurrence, and its occurrence in multiple settings.

The problem with accuracy, however, is obtaining measures of behavior that can legitimately be considered true values. In particular, when systematic direct observation by humans are used to collect data, developing an incontrovertible index against which to compare scores generated from such an observation system can prove difficult. When possible, some form of technological reproduction is typically used to generate the true values of the behavior. For example, sample sessions of behavior may be videotaped and studied carefully in order to determine the true values for the target behavior. The ability to replay the tape as many times as necessary or to use stop action and reverse options helps avoid error, although it can never be proven that error is completely absent. In this manner, observed behavioral values can be compared to the true value template and accuracy assessed.

The problems in assessing accuracy are obvious. First, the practicality and feasibility of obtaining true values or an incontrovertible index of behavior are low. Simply, most behavioral assessors and many researchers are not equipped with the type of technological assistance that is required to capture the true essence of the target behavior. Second, the issue that one can never fully know the true value behavior because it is impossible to separate what we know from how we know it is ever present. Therefore, accuracy is highly time and context dependent and only represents point estimates of reliable measurement. For these reasons, accuracy is rarely considered or reported in the behavioral assessment literature.

In addition to accuracy, a second commonly used proxy for reliability is that of interobserver agreement. Here, concurrent observations are made independently by two observers and the degree of association between the observers is assessed. It is important to note, however, that interobserver agreement provides no information with respect to accuracy and reliability. Reliable or accurate observations are those that demonstrate a consistent relationship to the behavior as it actually occurs; interobserver agreement provides no such information. Simply, the fact that two observers agree on the number of times a target behavior occurred says nothing about whether their observations were accurate or reliable. For example, two observers may both note the occurrence of 10 instances of the target behavior. The level of their agreement would appear to be 100%. However, if 20 instances of the behavior actually

occurred, and observer A witnessed only the odd-numbered occurrences and observer B noted only the even-numbered occurrences, their actual agreement would be 0%. As such, although complete agreement between two observers might seem like a comforting piece of information, the only conclusion that can be made is that their total values (not the actual behaviors) for the session were in agreement.

As can be seen, both interobserver agreement and accuracy involve comparing an observer's data to some other source. However, they differ considerably in the extent to which the source of comparisons can be entrusted to reflect the true behavior as it actually occurred. Accuracy, which makes a direct comparison to an incontrovertible index of the real behavior, is obviously the more reliable (and valid) of the two; however, it is difficult to estimate in applied research settings. Interobserver agreement, although easier to acquire and often used interchangeably with reliability, suffers from an inability to demonstrate accuracy and makes obvious the problems with using consensus as a replacement for reliability.

8. SUMMARY

Behavioral observation is one of the most widely used assessment strategies in schools. Given its flexibility and ease of use, behavioral observation procedures can be used to collect a range of data that provide helpful information and are useful for making a variety of psychoeducational decisions. Because of its direct nature, behavioral observation is particularly well suited for everyday life settings and can provide a systematic record of behavior that can be used in preliminary evaluation, intervention planning and design, the documentation of changes over time, and as part of a multi-method–multisource evaluation that integrates other forms of assessment (e.g., interviews and rating scales) and sources (e.g., teachers, parents, and children).

See Also the Following Articles

Effective Classroom Instruction ■ Emotional and Behavioral Problems, Students with ■ School Discipline and Behavior Management ■ School Environments

Further Reading

Alessi, G. (1988). Direct observation methods for emotional/behavioral problems. In E. S. Shapiro, & T. R. Kratochwill (Eds.), *Behavior assessment in schools: Conceptual foundations and practical applications* (pp. 14–75). New York: Guilford.

Hintze, J. M., Volpe, R. J., & Shapiro, E. S. (2002). Best practices in the systematic direct observation of student behavior. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology—IV* (Vol. 2, pp. 993–1006). Washington, DC: National Association of School Psychologists.

Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children & adolescents.* Mahwah, NJ: Erlbaum.

Skinner, C. H., Dittmer, K. I., & Howell, L. A. (2000). Direct observation in school settings: Theoretical issues. In E. S. Shairo, & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed., pp. 19–45). New York: Guilford.

Boredom

William K. Balzer, Patricia C. Smith and Jennifer L. Burnfield

Bowling Green State University, Bowling Green, Ohio, USA

GLOSSARY

affect circumplex A two-dimensional model of affect, with orthogonal dimensions of pleasure–displeasure and high–low activation (arousal).

flow Experiences in which people are highly engaged and focused on what they are doing. When in a flow state, one can complete tasks well and with ease.

monotony Perceived repetitiveness or tedium of a task.

output curve Plot of the rate of output produced during a specified time interval (e.g., a workday).

vigilance The ability to maintain attention and respond to stimuli over an extended period of time.

Boredom is an unpleasant, short-lived, affective state in which a person feels a pervasive lack of interest in, and difficulty attending to, a task or activity. Despite the absence of a single model or theory to explain boredom at work, research has made some progress on identifying its causes and a wide range of interventions to prevent or alleviate its occurrence.

1. INTRODUCTION

Interest in boredom at work can be traced to the early 1900s. In 1913, Munsterberg, the father of applied psychology, identified boredom as a valuable area of inquiry in the workplace. The first scientific study, "The Physical and Mental Effects of Monotony in Modern Industry," was conducted more than a decade later by Davies in 1926. Since these early studies, the research on boredom has been sporadic and fragmented, resulting in a nascent understanding of a construct that is thought to have significant performance outcomes and personal consequences for employers and employees. Smith's 1981 article succinctly summarizes the then—and current—state of research on boredom at work: "The amount of research devoted to the topic of boredom by psychologists and psychiatrists is astonishingly small when compared to literary treatments and to the acknowledged importance of [boredom at work]" (p. 338). Although the focus of this article is on boredom at work, it is important to recognize that other areas of applied psychology have also investigated the causes and consequences of boredom, including education, occupational health, ergonomics, military psychology, mental health, and interpersonal relationships.

In fact, the recent research on boredom has not been conducted in the domain of work. Much of the recent work has been in the areas of clinical psychology and personality.

2. DEFINITION OF BOREDOM

A single clear definition of boredom has been elusive. Boredom has often been used interchangeably with other related but not identical constructs (monotony, repetitiveness, ennui, and tedium) or its observed or hypothesized phenomenological correlates (fatigue, irritation, depression, dissatisfaction, hopelessness, and stress). Others have approached boredom by focusing on methods of alleviating boredom: novelty, interest, complexity, curiosity, enrichment, and "flow" (i.e., experiences in which people are highly engaged, focused, and able to do well). Progress has been made toward a single definition of boredom, complicated by the fact that differing models or theories of boredom (e.g., psychological, psychophysiological, and motivational) necessarily affect its definition. Commonalities across definitions lead to the working definition of boredom used here, as proposed by Fisher in 1993: "Boredom is an unpleasant, transient, affective state in which an individual feels a pervasive lack of interest in and difficulty concentrating on the current activity." By definition, boredom is viewed as a negative reaction in response to a personal assessment of a work situation that is more short-lived than attitudes such as job satisfaction. The relation of boredom to other affective constructs in a larger nomological net has been acknowledged. In the framework provided by the "affect circumplex," "boredom" is located on the negative ends of the orthogonal dimensions of pleasantness and activation, along with "tired" and "drowsy."

3. MODELS OF BOREDOM

Despite 90 years of investigation, there is no single, agreed upon model of boredom at work. Smith's 1981 research provides a good historical overview of the development of boredom at work, including some models of the causes and consequences of boredom: conflict models, psychoanalytic models, signal detection models, cognitive models, and drive models. Four models serve here as exemplars of the existing approaches to the understanding of boredom at work:

the psychodynamic model, the cultural normative model, the psychophysical model, and the cognitive–affect model.

3.1. Psychodynamic Model

Sigmund Freud's psychodynamic approach to the understanding of underlying psychological forces has been applied to explain boredom at work. Boredom at work is thought to reflect an inability to identify and ascribe emotional meaning to one's work. In this view, workers avoid or deny finding meaning in their work because they expect that individual work is not meant or designed to provide personal meaning. The meaninglessness of their work is experienced as boredom, which may be hidden from awareness and understanding through repression, denial, and defense mechanisms, resulting in tension (between what a person wants and what a person gets from work) that manifests itself in psychosomatic complaints, pathology, and so forth.

3.2. Cultural Normative Model

This model proposes that prior socialization makes employees highly dependent on authority figures to ascribe meaning to their work. Workplace norms develop that lead workers to accept without question boredom at work. In this workplace, boredom is denied and suppressed. Individual workers are not free to share or discuss their perceptions or emotions due to these normative pressures. Instead, they are relegated to chronically boring and meaningless jobs.

3.3. Psychophysical Model

The psychophysical model of boredom at work holds that the human nervous system is an active searcher, rather than a passive recipient, of environmental stimulation. Boredom results from exposure to monotonous or repetitious sensory stimulation. Habituation or reduced arousal to the task will result unless the worker voluntarily exerts additional attention and effort to the task. This increased overall arousal offsets the declining sensory stimulation provided by the monotonous task while maintaining a level of arousal adequate to satisfactorily perform the task. Overall, a psychophysical model suggests that if individuals cannot compensate for inadequate arousal from a task by increasing their level of effort, boredom will result.

3.4. Cognitive–Affect Model

The cognitive–affect model proposed by Hill and Perkins in 1985 posits that boredom is a psychological (as opposed to psychophysical) phenomenon. If the sensory stimulation from work fails to meet the needs or expectations of the individual employee, and the employee is unable to supplement the task through additional or alternative stimulation, frustration (negative affect) will result, leading to the overall experience of boredom. In this model, it is the subjective interpretation of the cognitive components of the task that is critical, rather than the actual sensory stimulation. If the task is perceived as monotonous and unable to meet the instrumental needs of the employee engaged in the task, it will lead to an affective disturbance and the experience of boredom.

4. MEASURING BOREDOM

Early assessments of boredom in the workplace were based primarily on the observation of workers, unstructured self-reports, or the inference of boredom based on patterns of responses to related items on structured surveys. For instance, early studies used subjective perceptions of output as a criterion of boredom. In this method, employees indicate the time of the day in which they think they work better and have higher production rates while also indicating when they felt bored. Studies of production workers found higher levels of boredom in the morning and attributed the increased effort and interest as the day progressed to the impending goal of reaching the end of the workday. This pattern of findings supports Wyatt's 1929 rich metaphor for boredom at work—a long hike:

> When activities provide little or no interest, the prospects are depressing and discouraging, and the effect seems to bear some relation to the magnitude of the task to be accomplished. To a person who dislikes walking, the prospects of a 20-mile tramp over uninteresting country may be almost painful, and although he starts with a certain amount of forced effort and enthusiasm, this soon disappears and leaves him in a darkened mood. At this stage the magnitude of the task may be appalling, but as the time for lunch and rest draws near interest is awakened and depression recedes. The distance still to be traversed now shrinks considerably, and bears little resemblance to the endless trail conceived a few hours before. A new swing

and liveliness characterize (sic) his step as he again takes to the road, and although these may diminish after a few more miles, they reappear with increased intensity as he enters the lap which separates him from his goal. (p. 162)

Several researchers sought an objective index of boredom in the curve of piece-rate output. This curve is a plot of the total number of pieces completed for some fixed period of time (e.g., 1 h intervals) throughout the day. Work output showed U-shaped curves with a sag in the middle of the morning and again in the middle of the afternoon. The declines in performance were attributed to slower rates of performance when machine operators were bored, with higher rates at the beginning of work and in anticipation of the end of the workday. However, when researchers actually observed and timed workers, the drop in production proved to represent voluntary rest pauses, including rest room and smoking breaks, which may be needed or scheduled at predictable times throughout the workday. Output rate was not slower when the operators were actually working at their machines. Overall, although the hiking metaphor may explain what is happening psychologically, observations of workers that experimentally control for employees who leave their workstations (e.g., a smoking break) do not necessarily support this assumed pattern of motivation and performance.

Following the era in which boredom in the workplace was measured through unstructured employee self-reports of boredom or objective measures of work output over time, a number of self-report rating instruments have been developed. For example, Lee's 1986 Job Boredom Scale measures boredom through employee responses to 17 items concerning satisfaction, interest, and connectedness to the job. Measures also exist that assess workers' susceptibility or proneness to experience boredom. The most common measures are Farmer and Sundberg's 1986 Boredom Proneness Scale and Zuckerman, Eysenck, and Eysenck's 1978 Sensation Seeking Scale (e.g., measuring aversion to repetition and dull people). Unfortunately, there is no single source reference that provides descriptions and psychometric evaluations (e.g., does the instrument measure boredom at work in a consistent and conceptually appropriate manner?) of available measures. College and university research libraries and Internet search engines may be the best sources for finding well-constructed measures for predicting or assessing boredom at work.

5. CAUSES AND CORRELATES OF BOREDOM

Despite the absence of a single unified model or theory to explain boredom at work, the state of research has progressed on identifying its causes. The hypothesized causes of boredom can be grouped into three general categories: the task, the work environment, and the person.

5.1. Task as a Cause of Boredom

The traditional view of boredom at work assumes that boredom resides in the task itself. Specifically, tasks that provide low stimulation (i.e., are repetitive and monotonous) have been thought to result in boredom. Fisher's 1993 article offers a conceptual framework of work tasks that may lead to boredom: quantitative underload (jobs that provide the worker with nothing to do or jobs with a high degree of workload variability, e.g., periods of inactivity after high levels of activity), qualitative underload (jobs that are simple, repetitive, limited in mental demand, unchallenging, or do not use the individual's skills), and qualitative overload (jobs that are too challenging, difficult, or incomprehensible).

5.2. Work Environment as a Cause of Boredom

The work environment may be seen as a lens through which feelings of boredom are reduced or intensified. The mere presence of coworkers can enhance arousal directly through conversation or interdependent work. However, dull, unhappy, or uncommunicative colleagues may actually contribute to feelings of boredom. From a social influence perspective, fellow employees may magnify or reduce boredom through statements or actions that make work seem challenging and interesting or boring and repetitive. Organizational policies and practices may also contribute to the presence or absence of boredom. Rules that prohibit talking, limit work breaks, or prescribe precise work procedure may be so constraining as to induce feelings of boredom.

5.3. Person as a Cause of Boredom

Individual differences among employees may also contribute directly to boredom at work. Research suggests that males tend to be more prone to boredom than females, particularly when there is a perceived lack of stimulation. However, this gender finding may be culture-specific. Individuals with high cognitive ability or capacity may be more likely to experience boredom on tasks that underutilize their abilities. Conversely, the more intelligent worker may perceive work differently and find a higher degree of variety in a seemingly homogeneous and repetitive task. Personality differences may also contribute to feelings of boredom. Extraverted people may require more stimulation to keep them aroused and engaged at work. The expectations and alternatives available to the person may also affect the perception of interest versus boredom. If a job is the only available job without the hope of other prospects, it may be perceived as less boring. The psychiatric literature has also identified pathologically bored individuals who are unable to perceive the stimulation inherent in work activities or are unable to focus their attention at work as do their better adjusted colleagues.

It is also probable that the task, environment, or person characteristics that cause feelings of boredom interact to accentuate or attenuate these feelings. For example, an employee with limited cognitive capacity performing a challenging and demanding task (i.e., qualitative overload) may experience strong feelings of boredom, whereas an employee with a greater intellectual capacity would not. Conversely, the high-capacity employee may feel more bored on a simpler task (i.e., qualitative underload) than would a person of more limited capacity. As another example, employees who do not perceive any value or instrumentality in their participation at work are likely to feel more bored regardless of whether their work tasks are repetitive or challenging. Finally, individuals who have more complex approaches to the world (i.e., high in schema complexity) may be able to perceive more nuances and details in their work, making them more engaging and less boring. They also may be more likely to engage in goal setting, enhancing the meaning and motivation from their work. Overall, the causes of boredom are likely to be a combination of the direct and interactive impact of task, environment, and person characteristics.

6. CONSEQUENCES OF BOREDOM AT WORK

Boredom at work has negative consequences for workers, their organizations, and society as a whole. Research

suggests that the predicted or demonstrated effects of boredom for individual employees include work stress, job dissatisfaction, physical and mental health problems, workplace hostility, increased risk taking or thrill seeking, horseplay and sabotage, alcohol and drug problems, absenteeism, work-related injuries among adolescent employees, and poor performance. There is also evidence that boredom leads to the perception that time passes more slowly, although individuals' objective estimates of time passed seemed to be unaffected by high levels of boredom. The consequences of boredom for businesses and organizations include higher employment costs (due to employee turnover, absences, and health care costs), performance problems (work quality and work quantity), and reduced organizational effectiveness (including profitability). Finally, boredom at work may also have consequences for society as a whole, including lost productivity, reduced quality of work life, and reduced consumer safety. Numerous studies of vigilance—the ability to maintain attention and respond to stimuli over an extended period of time—reveal that an observer's ability to detect critical signals declines over time. This may be exhibited by a reduction in accuracy or increased response time. High levels of boredom experienced in vigilance tasks have been associated with stress and potentially more serious outcomes. Especially disconcerting is the case of a commercial airliner that overflew its destination by more than 100 miles before the sleeping flight crew was awakened by air traffic controllers. Clearly, boredom at work has significant consequences.

7. PREVENTION AND ALLEVIATION OF BOREDOM

Repetitive assembly, continuous manual control, and inspection and monitoring tasks may be most likely to lead to boredom at work. A wide range of interventions have been suggested to prevent or alleviate boredom at work. One would suspect that the intervention would be most effective if it corresponded to the cause of boredom (e.g., if boredom is caused by the task, the intervention should also focus on changing the task or the perception of the task). Consistent with the framework for categorizing the causes of boredom at work, interventions to prevent or alleviate workplace boredom are therefore organized around the task, the work environment, and the person.

7.1. Task Interventions to Prevent and/or Alleviate Boredom at Work

Although some work tasks may be inherently boring, interventions have been identified that can help minimize boredom on even these tasks. One category of intervention is to design work that is stimulating and psychologically rewarding rather than monotonous and fractionated. Morgeson and Campion described a number of conceptual models for redesigning work to reduce the effects of stressors such as workload and repetitive work. Such interventions can be focused on the motivational aspect of the task (e.g., job enrichment to enhance task meaningfulness) or a reduction in exposure to the stressor that induces boredom or its effects (e.g., ergonomic changes in the work environment to reduce the effects of repetitive work).

In addition, for repetitive tasks, grouping output into recognizable "lots" or batches and providing a system for tracking the number of lots completed may be a useful intervention. Job rotation, or changing work assignments throughout the day, may help employees stay alert and challenged at work. Finally, the use of performance goals (either self-set goals or the acceptance of company-set goals) and feedback on goal success can help prevent or alleviate boredom at work.

7.2. Work Environment Interventions to Prevent and/or Alleviate Boredom at Work

Employee selection and placement systems are an underused means for preventing boredom. Matching individuals to particular jobs may be particularly effective given differences in individual perceptions and reactions to different work tasks. The use of performance-based reward systems, including piece rate payment systems, may also create an incentive that increases employee effort and may prevent or reduce boredom for those who are prone to boredom. However, workers with a low proneness to boredom may benefit from jobs and work environments that emphasize intrinsic rewards such as job enrichment. Companies can also provide rest periods, which have been found to neutralize the unpleasant and unfavorable effects of repeated activity. Providing work that is employee paced rather than externally paced (e.g., a conveyor system) can allow the employees either individually or as a group to adapt their workload to cyclical levels of attention and effort throughout the day. Communicating the importance of the employee's work

contribution to the overall effectiveness of the work unit or company can increase interest in and attention to work. Finally, designing the workplace so that it allows for conversations and social interactions among employees can help prevent or alleviate boredom at work.

7.3. Person Interventions to Prevent and/or Alleviate Boredom at Work

Employees can take steps to reduce boredom at work. Fisher's 1998 article demonstrated that brief interruptions by both internal (e.g., mind wandering) and external (e.g., mobile phones and e-mail) sources may alleviate boredom, depending on the level of attention required to perform the task. In terms of internal mechanisms, forcing oneself to refocus attention or effort on the task can help minimize boredom, particularly at the outset of unstimulating tasks or in constraining work environments. Employees can also seek additional stimulation in an unstimulating work environment by playing mental games, assisting coworkers, and so forth. For well-learned repetitive tasks, employees might use daydreams to escape a tedious and monotonous task. Individual employees may also cope with boredom by learning to integrate their rhythm with that of machines, allowing them to accomplish their tasks through more automatic attentional processes. External interruptions have also been found to reduce boredom on simple low-attention tasks, perhaps by providing variety and additional stimulation to a monotonous work environment. However, such interruptions may be problematic in other situations or for other outcomes.

8. CONCLUSION

Given the individual, organizational, and societal consequences of boredom at work, it is surprising that more attention and effort have not been devoted to theory development, psychological measurement, and scientific investigation. The available research, spanning 90 years of work by practitioners and scientists, has provided a good foundation for expanding our knowledge and understanding of boredom at work. How will current and future changes in the workplace affect boredom at work? The growing use of robots and computers to complete repetitive and monotonous tasks could potentially reduce boredom at work. The migration of simple manufacturing and assembly jobs to other countries may, on average, reduce the level of boredom in the U.S. workplace—if

replacement jobs offer more stimulating tasks or work environments. Changes in the workplace that enhance workers' control over their jobs (i.e., meaningful assignments, goal setting, and feedback) and meet their social needs at work may reduce the level of boredom for future employees. Additional research is needed to ensure the prevention and treatment of boredom in today's workplace and the changing workplace of the future.

See Also the Following Articles

Competence at Work ■ Cooperation at Work ■ Work Adjustment ■ Work Environments ■ Work Motivation ■ Work Role, Values Sought in the ■ Work Teams

Further Reading

Davies, A. H. (1926). The physical and mental effect of monotony in modern industry. *British Medical Journal*, 2, 472–479.

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness: The development and correlates of a new scale. *Journal of Personality Assessment*, 50(1), 4–17.

Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations*, 46(3), 395–417.

Fisher, C. D. (1998). Effects of external and internal disruptions on boredom at work: Two studies. *Journal of Organizational Behavior*, 19, 503–522.

Gemmill, G., & Oakley, J. (1992). The meaning of boredom in organizational life. *Group & Organization Management*, 17(4), 358–369.

Hill, A. B., & Perkins, R. E. (1985). Towards a model of boredom. *British Journal of Psychology*, 76, 235–240.

Kass, S. J., Vodanovich, S. J., & Callender, A. (2001). State-trait boredom: Relationship to absenteeism, tenure, and job satisfaction. *Journal of Business and Psychology*, 16, 317–327.

Lee, T. W. (1986). Toward the development and validation of a measure of job boredom. *Manhattan College Journal of Business*, 15, 22–28.

O'Hanlon, J. F. (1981). Boredom: Practical consequences and a theory. *Acta Psychologia*, 49, 53–82.

Morgeson, F. P., & Campion, M. A. (2003). Work design. *Handbook of Psychology: Industrial and Organizational Psychology*, 12, 423–452.

Smith, R. P. (1981). Boredom: A review. *Human Factors*, 23(3), 329–340.

Sommers, J., & Vodanovich, S. J. (2000). Boredom proneness: Its relationship to psychological- and physical-health symptoms. *Journal of Clinical Psychology*, 56, 149–155.

Wyatt, S. (1929). Boredom in industry. *Personnel Journal*, 8(3), 161–171.

Zuckerman, M., Eysenck, S., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology*, 46, 139–149.

Bullying and Abuse on School Campuses

Michael J. Furlong, Alicia C. Soliz, Jenne M. Simental, and Jennifer L. Greif

University of California, Santa Barbara, California, USA

Most adults can remember bullies and victims of bullying from when they were in school. The impact of these events does not stop when students graduate; these events can have long-lasting effects on bullies, victims, and even students who only watched bullying occur. The term bullying refers to a specific type of peer aggression that is intentional, repeated, and involves an imbalance of power between the bully and the victim. Suggestions for intervention approaches at the individual, classroom, and schoolwide level are provided in this article.

GLOSSARY

bullying Aggression that is intentional and repeated over time, and there is an imbalance of power between the perpetrator and the victim. Bullying is a specific type of peer victimization that is categorized by these components, which indicate an ongoing relationship between the bully and the victim.

direct bullying When a bully is overtly aggressive toward a victim (e.g., repeatedly hits, teases, or steals from the victim).

indirect bullying When a bully expresses aggression toward a victim through a third party (e.g., spreads rumors and gossips). Although the bully and the victim do not actually interact with each other in instances of indirect bullying, this aggression still has an impact on the well-being of the victim.

1. INTRODUCTION

Bullying in school settings was once thought of as a normal, transient part of growing up. It was considered to have little lasting impact, and the best advice to the victim was to fight back or turn the other cheek. Beginning in the late 1970s and through the 1980s, pioneering research began to shed light on this behavior and to dispel common myths. Expanded studies during the 1990s started to examine bullying as a serious form of aggression and considered the possible impacts of this experience on the bully, his or her victim, and even bystanders. The results of these investigations have shown that not only is bullying a common occurrence but also it has lasting effects on all involved.

The pioneer work on bullying was conducted in the mid-1980s by Dan Olweus, a Norwegian psychologist commissioned to conduct a large-scale study on the topic. Norway's interest in bullying was driven by a string of suicides committed by adolescent boys who had been victims of severe bullying. These efforts were followed by a number of other investigations conducted by researchers in England (Smith), Canada (Peppler), and Australia (Slee, Rigby, and Griffiths). Detailed studies in the United States did not begin to focus seriously on bullying until later, primarily in response to several deadly school shootings that were perpetrated by students who were thought, at least in part, to be seeking revenge for past abuse by school bullies.

2. WHAT IS BULLYING?

Bullying is defined as aggression between peers that has three essential elements: It is intentional, it is repeated over time, and there is an imbalance of power between the bully and the victim. This definition (developed by Olweus) distinguishes bullying from other types of peer aggression, such as one-time fights or friendly teasing between friends. It is these three defining characteristics that make bullying potentially so damaging. Victims are intimidated or hurt repeatedly by someone who is more powerful; consequently, they may learn to accept the victim role and come to distrust others. Ongoing victimization can have a serious impact on the development and well-being of children who are bullied. In addition, bullies who abuse their power (e.g., physical strength and popularity) to single out and intimidate peers potentially become involved in a pattern of using this social control strategy throughout their lives. As such, bullying is considered an exploitative relationship between peers rather than a one-time event.

Despite relative consensus among researchers as to what constitutes bullying, this clarity does not always extend into practice. Many of the most commonly used measures of bullying do not specify that, by definition, bullying is a series of intentional, repeated acts of aggression that occur in peer relationships in which there is an imbalance of power. The omission of a precise definition in research is problematic because, without these specifications, studies may overreport the prevalence of bullying when students are asked about general aggression. In addition, when educators and others think of bullying as a specific event, rather than a series of encounters, interventions may be more punitive and less focused on changing the process of interactions between peers.

In the past, bullying was considered to be physical acts of aggression that primarily occurred between boys. It is now recognized that bullying can take many forms, including ongoing physical abuse (e.g., hitting and pushing), verbal abuse (e.g., taunting and name-calling), social manipulation (e.g., rumor spreading and purposeful exclusion from activities), and attacks on property (e.g., breaking or stealing belongings). One common distinction is between direct bullying (i.e., physical aggression or chronic teasing that occur overtly between the bully and the victim) and indirect bullying (i.e., gossip or rumors targeting the victim that occur between the bully and a third party). Recognition that bullying can include these additional behaviors has broadened the definition and it is now more inclusive of types of aggression that occurs more frequently between girls (e.g., gossip and exclusion). Regardless of its specific form, all bullying involves the repeated abuse of a student by another more powerful student who wants to harm the victim. The bully, in fact, comes to enjoy exercising power over his or her selected target. It is a myth that bullies abuse others to compensate for low self-esteem. In fact, they often have positive self-esteem, may be popular with their peers, and inappropriately come to enjoy having power over others.

3. HOW COMMON IS BULLYING ON SCHOOL CAMPUSES?

The incidence of bullying varies from country to country and school to school within countries. However, it is certain that bullying occurs at all comprehensive schools across all grade levels. Students consistently report that most bullying incidents take place away from an adult presence at the school, including on school playgrounds, during or between class, and walking to and from school. In addition, bullying is often unreported by the parties involved, primarily because the bullies wish to avoid punishment and the victims wish to avoid further aggravating the bully or prompting retaliation—bullies intimidate their victims. Current research practices for collecting information utilize a combination of teacher and/or parent reports as well as student self-reports and student identification or nomination. Because researchers use various methods, a definitive estimate of the prevalence of bullying in schools is difficult to obtain. It is also probable that the available statistics underestimate the extent and scope of this problem.

A 1998 survey of 15,686 students in the United States conducted by the World Health Organization found that 16% of boys and 11% of girls felt bullied, with 23% of boys and 11% of girls admitting to bullying other students. Research has also found an association between the frequency of short-term and long-term victimization. For example, in a 2003 study of Norwegian victims of bullying, Solberg and Olweus found that bullying that persisted over a period of 1 month or more was more frequently reported than short-term bullying events. In other words, bullying is often a chronic experience for its victims.

There are some developmental differences in bullying. A 1993 study by Whitney and Smith found that the occurrence of bullying decreases as students get older, with the highest rates occurring in the upper elementary and middle schools and the lowest rates in secondary schools. This is suspected to be due, in part, to increased social and physical maturity. By the time a child reaches secondary school, there is a decreased presence of older, more powerful peers who can single out and pick on smaller, weaker peers.

In addition to age differences, there are also gender differences in bullying trends. As mentioned previously, boys are more likely to report both bullying and victimization; however, girls are more likely than boys to experience indirect bullying as opposed to direct bullying. Although the consequences of social rejection and isolation associated with indirect bullying may not be immediately obvious, it may produce negative long-term effects on the victim's social development. Finally, female victims are more likely to tell a female peer about their experience, whereas males are more likely to turn to a trusted adult.

4. WHO ARE THE BULLIES?

Despite the increased awareness of female bullying, several researchers have found that most bullies tend to be boys or groups of boys. The results of studies designed to identify the etiology of bullying behavior have been mixed. Some experts suggest that bullies act to exert power over others with the primary purpose of gaining more power. Alternatively, other experts draw upon a social learning model in which bullying is a reaction to being bullied by other children or a reaction to a challenging home environment with associated authoritarian discipline experiences.

In cases in which bullies appear to be the popular children with no history of victimization, the bully behavior may be somewhat puzzling. In these cases, experts have suggested that bullying is a reaction to

some underlying need for power. Bullies are rewarded for exerting power over others by gaining greater power, which reinforces this negative behavior.

Alternatively, the social learning perspective suggests that children who have suffered bullying may begin to bully others less powerful than themselves either as a form of retaliation or as a way of gaining some sense of empowerment. These children have been categorized as bully/victims. Similar to these bully/victims, children from abusive family environments may experience bullying at the hands of family members and, in turn, become bullies themselves. For example, Roberts and Morotti suggest that bullying behavior stems from abusive, chaotic home environments in which families bully the child, who in turn bullies a peer (i.e., a "kick-the-dog" phenomenon). In this type of environment a child learns to normalize and eventually utilize negative social interactions, such as bullying. This idea is supported by the identification of common characteristics of children known to suffer from abusive family experiences. For example, children who bully often display behaviors indicating low social competence, antisocial behavior, a lack of empathy, and high levels of aggression (physical, verbal, and relational). In addition, when interviewed, bullies often describe home environments that are more negative and authoritarian in style than those of children without histories of bullying.

5. WHO ARE THE VICTIMS?

Identified victims of bullying share many similar characteristics. In one pattern, victims are classified as either passive (never behaving aggressively) or highly aggressive (often provoking peers). Most victims are children who spend more time alone than typical children, often falling into the socially rejected category when rated by their peers. They often have poor or less than desirable physical characteristics (e.g., smaller in height, weight, and strength), are younger than the majority of their peers, display low social competence, and emotional maladjustment including depression. In general, victims of bullying are those youth whose physical, psychological, or social status is diminished compared to that of school bullies, thus making them possible targets of repeated aggression.

6. WHAT ARE THE EFFECTS?

The commonly cited short-term effects of bullying to the victim include risks of diminished self-esteem and

depressed mood. Several physical symptoms, such as stomachaches and headaches, difficulty concentrating, and disruption of sleeping and eating habits, have been present among victims of bullying. Long-term effects can include depression, social anxiety, increased levels of aggression, lowered self-esteem, increased risk of suicide, decreased coping skills, and the presence of posttraumatic stress disorder symptoms. As a result of being traumatized by incidents of bullying, victims may carry these negative consequences into adulthood, leading to decreased levels of societal functioning.

Less well known are the effects of bullying on the bully himself or herself. It appears that the long-term effects include risks of school failure; further antisocial, criminal, and delinquent behavior; and an increase in depression, suicidal ideation, and global negative self-evaluations. In 1993, Olweus found that of the Norwegian children identified as bullies in grades 6–9, 60% were convicted of at least one crime by the age of 24. Surprisingly, of this 60%, 35–40% had three or more convictions. In addition, incidents of drug abuse in conjunction with depression and suicidal tendencies have often been high among individuals identified as bullies, especially in comparison to individuals with no history of bullying. Finally, bullying is a behavior that, once engaged in, becomes a pattern that is difficult to stop. The bully may become accustomed to and rewarded by getting what he or she wants through asserting power over other individuals. This behavior may continue throughout life and affect social interactions and interpersonal relationships. In short, someone who is abusive in one aspect of his or her life tends to be abusive in various other aspects of life, which can lead to an aggressive and typically frustrating pattern of social behavior.

7. WHAT CAN BE DONE?

Bully intervention programs implemented at schools seek to eliminate existing bully problems, prevent the development of new bully problems, and achieve better peer relations at school. Other goals of bully intervention programs are to create a positive school climate and increase caring behaviors toward bully victims. These goals are accomplished at the schoolwide level, the classroom level, and the individual level.

7.1. Schoolwide Responses

Before implementing a bully prevention program, the following must be considered:

1. Staff training: What type of staff training is required? Who will provide training? How will training occur?

2. Program funding: How much funding is available? How will funds be allocated?

3. Time commitments: How much time will be required by the trained individuals outside of normal duties? How often will the program be implemented and for how long?

4. Clear behavioral objectives: What will be expected of students after the intervention? Will the school rules be changed?

5. Alternative programs: Which programs will be available for students who are not affected by the intervention?

6. Program leadership: Who will take the lead? How will staff be chosen to participate? What qualifications and training are necessary?

In order to equip schools with appropriate tools to deal with bullying, training should include an awareness component through which the school staff is informed about the extent of bullying on campus. Staff should also be asked to examine their own beliefs about bullying. Training on bully prevention and intervention should include specific strategies as well as an overarching conceptual understanding of the nature of bullying as described in previous sections.

Because bullying often happens in unsupervised contexts, such as on the playground and in the lunchroom, where paraprofessionals and community volunteers often provide supervision, bus drivers, maintenance staff, and cafeteria workers should also receive training. Support staff should be aware of the school code and the school's general bully response procedures.

Bully intervention programs at the school level often restructure the school environment to remove the consequences and negative impacts of bullying. Comprehensive programs typically create specific school-site committees that coordinate aspects of the program, administer questionnaires to determine the nature and extent of the bully problem, hold meetings to raise community awareness, improve supervision and the outdoor environment, and involve parents.

7.2. Classroom-Level Responses

As part of an overall schoolwide plan, teachers implement many bully prevention programs within their classrooms. These programs ideally include a set of clear classroom rules and routines (preferably shared

across classrooms at the same grade level), comprehensive academic lesson preparation, ample classroom structure, opportunities for students to develop empathy for one another, a warm classroom environment, cooperative instructional groups, and shared responsibilities. In addition to basic preparation and creating a positive classroom environment, teachers should never overlook bullying incidents because in doing so they may inadvertently contribute to a broader "culture of bullying" on the school campus. It is important to create conditions that encourage both victims and bystanders to report bullying when it occurs. It is also important to implement programs that address the unique needs and provide assistance to all involved—bullies, victims, and bystanders. Many teachers also implement programs to develop student skills for resolving conflict in order to provide opportunities for "in-the-moment" social skill learning.

7.3. Individual-Level Responses

It is important to determine the function that bullying serves in specific incidents because this will influence the intervention approach. Most bullies know that what they are doing is wrong, but they often have limited empathy for the victim and may have a personal need to dominate others. Merely punishing the bully (by using power over him or her) may only reinforce the notion that power is an effective social tool. Immediately following a bullying incident, a teacher or staff member should have a serious talk with the bully to determine what steps should be taken next. The talk should include a documentation of the incident, a message that bullying is not acceptable, a referral to the school's code of conduct, and a reminder that the bully's behavior will be closely monitored in the future. Individual responses to bullying also include continued one-on-one discussions, counseling, parent involvement, follow-up discussions, and, in extreme cases, the change of class or school for the bully or the victim.

In addition to talking to the bully, the victim should be interviewed as soon as possible, separately from the bully. It is important to obtain information about the duration and frequency of the bullying incident(s), if the victim has had similar experiences in the past, and what the victim has done to try to stop it. During this talk, the incident and nature of discussion should also be documented, the victim should receive information about the plan of action for the bully, and the victim should be taught how to report any future incidents. Assessing the social skills and status of the victim should

be carried out as part of an effort to provide social skills training to enhance his or her sense of personal efficacy to cope effectively with any future bullying.

Because peers play different roles in the bully process, it is important that they be considered as a possible resource in the intervention process. Training for students should also contain an awareness component in which they are taught about the nature of bullying specific to their school and the resources available to them. It is important that their consciousness is raised about the general processes of bullying and the individual's responsibility as a bystander. Training should focus on increasing empathy for the victims as well as provide peers with strategies to encourage them to withstand the pressures of their peer groups. Despite the common occurrence of bullying on school campuses, most students are not victims or perpetrators. They can have a powerful effect on the school climate as it relates to bullying. In part, bullying occurs on school campuses because it can. If bystanders recognize their important role in setting school standards, then its occurrence can be significantly reduced.

8. THE SPECIAL ROLE OF BYSTANDERS

Because peers are such an important part of the bullying process, they merit additional in-depth examination. By its nature, bullying is a social interaction that is embedded in a broader school context. The aggression that occurs between bullies and victims is rarely restricted in its impact to the bully–victim dyad. Rather, peers play an integral part in bullying and intervening. Olweus conceptualized student involvement in an acute bullying situation as being part of a continuum (or "bullying circle") (Fig. 1). At one end of the continuum are students who take an active part in bullying, support the bully, or follow the bully. At the other end of the continuum are students who defend the victim and dislike the bully's actions. Many students fit in roles that fall somewhere in between—supporting the bully but not taking an active part, supporting the victim but not helping in his or her defense, or as disengaged onlookers.

Simply by observing the bullying behavior, peers reinforce bullies who often gain social prestige through their aggression. In one study of school bullying, O'Connell, Peppler, and Craig found that when a higher number of peers were observing an incident of

The bully/bullies	Followers/henchmen	Supporters/ passive bullies	Passive supporter/ possible bullies
Start the bullying and take an active part	Taken an active part but do not start the bullying	Support the bullying but no active part	Like the bullying but no display of active support

The victim

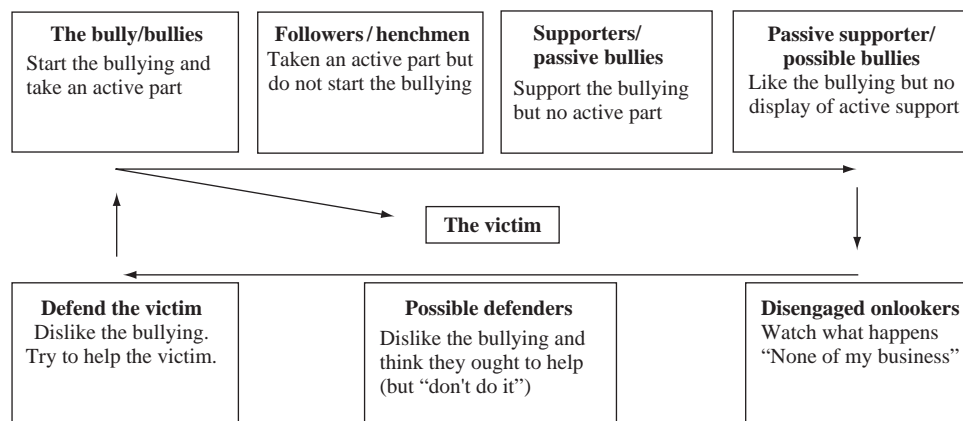| Defend the victim	Possible defenders	Disengaged onlookers
Dislike the bullying. Try to help the victim.	Dislike the bullying and think they ought to help (but "don't do it")	Watch what happens "None of my business"

FIGURE 1   Continuum of student reactions in a situation involving acute bullying. Adapted from Olweus (2001).

bullying, the duration of the incident was longer. The presence of peers served to encourage the bully by giving him or her more peer attention. At least for some bullying behavior, one of the primary reasons that bullying occurs is because the bully gains social prestige through this peer attention. When peers are asked for their opinions of bullies and victims, they usually say that the bully is more likeable than the victim (this is thought to be true because the popularity and aggression of bullies are more appealing than the perceived weakness of victims).

Although the influence of peers encourages bullies, the intervention of peers is also one of the most powerful ways to stop bullying. In one investigation, bullying stopped in three-fourths of cases when a peer intervened on behalf of the victim. Because adults are often unaware of instances of bullying, the impact of this peer intervention should not go underestimated. Nonetheless, although 41% of students said that they would like to help the victim, in reality only 25% intervened. The question of what leads peers to action or inaction in the face of a bullying situation is an important consideration that should be discussed openly in every school.

Students with higher social status are more likely to intervene and help the victim than students who are less popular and perhaps fear that they will also be bullied. Another reason for failing to help the victim may be related to the fact that many acts of bullying occur in front of a group. When several students together observe a peer being bullied, they may feel a lack of individual responsibility to help the victim (social psychologists call this the bystander effect). Even when students would like to intervene, they may not have the skills or know how to help the victim in a way that will be effective. These are topics that are addressed in several intervention programs that educate peers and involve the entire school community.

9. IMPLICATIONS

Research has identified the "culture of bullying" on school campuses. Students are less likely to report witnessing bullying or being personally victimized by a bully if they also perceive that in the past teachers at their school have not intervened to stop bullying. Thus, although many prevention and intervention efforts attempt to decrease the aggression of the bully and to enhance the social skills of victims, it is now recognized that if a school has a laissez-faire, inferred acceptance of, or a blind unawareness of bullying, then students are more likely to accept its occurrence as a "normal" part of the school experience. These developments in bullying research show that schools cannot afford to let students work out their conflicts. To do so may send a message that behaviors such as bullying are tacitly tolerated. Bullying is the largest violence/aggression problem in schools worldwide. As such, it is of the utmost importance that schools study their particular circumstances and take action to prevent its occurrence and to intervene when it occurs.

See Also the Following Articles

Aggression and Culture ■ Emotional and Behavioral Problems, Students with ■ School Discipline and Behavior Management ■ School Violence Prevention

Further Reading

Espelage, D. L., & Swearer, S. M. (2003). Research on school bullying and victimization: What have we learned and where do we need to go? *School Psychology Review, 32,* 365–383.

Juvonen, J., & Graham, S. (Eds.). (2001). *Peer harassment in school: The plight of the vulnerable and victimized.* New York: Guilford.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among U.S. youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, 285, 2094–2100.

O'Connell, P., Pepler, D., & Craig, W. (1999). Peer involvement in bullying: Insights and challenges for intervention. *Journal of Adolescence*, 22, 437–452.

Olweus, D. (2001). Peer harassment: A critical analysis and some important issues. In J. Juvonen, & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 3–20). New York: Guilford.

Olweus, D., Limber, S., & Mihalic, S. (1999). *Blueprints for violence prevention series: Book nine: Bullying Prevention Program (BPP)*. Boulder: University of Colorado, Center for the Prevention of Violence. (Information available at www.colorado.edu/cspv/blueprints).

Rigby, K. (2004). What it takes to stop bullying in schools: An examination of the rationale and effectiveness of school-based interventions. In M. J. Furlong, M. P. Bates, D. C. Smith, & P. Kingery (Eds.), *Appraisal and prediction of school violence: Methods, issues, and contexts*. Hauppauge, NY: Nova.

Roberts, W. B., Jr., & Morotti, A. A. (2000). The bully as victim: Understanding bully behaviors to increase the effectiveness of interventions in the bully–victim dyad. *Professional School Counseling*, 4, 148–155.

Smith, P. K., *et al.* (2002). Definitions of bullying: A comparison of terms used, and age and gender differences, in a fourteen-country international comparison. *Child Development*, 73(4), 1119–1133.

Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior*, 29, 239–268.

Unnever, J. D., & Cornell, D. G. (2003). The culture of bullying in middle school. *Journal of School Violence*, 2(2), 5–27.

Whitney, I., & Smith, P. K. (1993). A survey of the nature and extent of bullying in junior/middle and secondary schools. *Educational Research*, 35, 3–25.
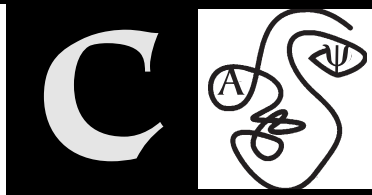
Career Counseling

Mary J. Heppner and P. Paul Heppner

University of Missouri, Columbia, Missouri, USA

GLOSSARY

career choice The act of choosing a career path; it is often first done during adolescence and early adulthood but is often modified as changes in the self, the world, or work change in the course of an individual's life.

contextual factors in career planning Factors in an individual's environment that may influence his or her career development; these factors might include the person's social class, race/ethnicity, family structure, and so forth.

self-assessment The process of gaining information about oneself (e.g., personality, values, skills, interests) so as to make an informed career decision; it can be conducted using inventories, using computerized career information systems, or having discussions with a career counselor.

world of work information Information that often includes job duties, education and training required, salaries and other benefits, and employment outlook; gathering world of work information is often an integral part of the career planning process.

Career counseling is ongoing face-to-face interaction performed by individuals who have specialized training in the field to assist people in obtaining a clear understanding of themselves (e.g., interests, skills, values, personality traits) and to obtain an equally clear picture of the world of work so as to make choices that lead to satisfying work lives. Career counselors help clients within the context of a psychological relationship with issues such as making career choices and adjustments, dealing with career transitions, overcoming career barriers, and optimizing clients' work lives across the life span. Career counselors are cognizant of the many contextual factors present in the lives of their clients and of the ways in which social and emotional issues interplay with career issues.

1. THE EVOLUTION OF THE DEFINITION

The current field of career counseling has been referred to by many different terms such as vocational counseling, career guidance, and vocational guidance. In fact, its definition has evolved to meet the changing context and needs of individuals within society. During the early writings about career counseling, it was described as a process of "matching men and jobs," reflecting the basic function of helping primarily male workers find jobs. As the field progressed, its purview was widened to involve helping both men and women with their career development more broadly and including issues such as decision making, career adjustment, and career

change. There have also been refinements of the definition, including an emphasis on career counseling being done over the life span as opposed to an activity primarily restricted to one's initial job choice. In addition, there has been growing acknowledgment of the intertwined nature of career counseling and social/emotional counseling, with some research indicating that career counseling may help individuals not only with their career adjustment but also with their personal and psychological adjustment. With the growing diversity in society, there has also been particular attention paid to the importance of the cultural context of an individual's life, including factors such as race/ethnicity, social class, age, ability level, sexual orientation, and religion. Thus, the definition and purview of career counseling continues to be dynamic, addressing the needs of a changing society.

2. HISTORY OF THE FIELD OF CAREER COUNSELING

The origins of career counseling, broadly defined, can be found during early Greek and Roman times, but the roots of the field as is practiced today can be traced to the last half of the 1800s and the early 1900s in the United States. This was a time when the origins of the human sciences were forming, particularly in experimental psychology, with the work of Wilhelm Wundt. It was also a time when "mental testing" was born, with James Cattell first using this term in an article in the journal *Mind.* Cattell was most interested in using mental tests to examine individual differences, particularly in measuring student achievement. Economically, this was a time of great turmoil and demographic change because the industrial revolution was under way and the country was moving from a primarily agrarian economy to an industrial-based economy. At this time, there was also an influx of immigrants, primarily from Europe, entering the country. Often, these individuals faced a host of social and employment problems in finding employment. Even once they secured jobs, they often faced unhealthy working conditions, overt discrimination, and physical and psychological abuse. There was a strong need for services to help these new workers find their place in the employment structure and to provide a range of social services and social advocacy on their behalf.

3. THE ROLE OF FRANK PARSONS

The need for a strong social advocate for these immigrants was found in the person of Frank Parsons, who is often referred to as the "Father of Career Development." Parsons developed what was termed the Breadwinners Institute under the Civic Service House in Boston. The institute was designed to provide a range of educational services for immigrants. It was in this context that Parsons developed a systematic way of helping individuals to find appropriate work that still has much influence on the way in which career counseling is conducted today. Parsons' book *Choosing a Vocation,* which outlined his decision-making process, was published in 1909 after his death. Parsons theorized that there were three broad decision-making factors: (a) a clear understanding of oneself, including one's aptitudes, abilities, interests, and limitations; (b) a knowledge of the requirements, advantages, disadvantages, and prospects of jobs; and (c) the ability to reason regarding the relation of these two sets of facts. These three factors have had an enormous impact on how career counseling has been practiced during the century since it was originally published.

4. THEORIES OF CAREER DEVELOPMENT

Career development theories provide foundational knowledge from which to draw useful concepts to explain behavior. They offer a framework within which individual career behavior can be explained and hypotheses can be formed about the possible meaning of behavior. As such, they help career counselors to identify and understand clients' goals and problems within a theoretical framework. Since the inception of the field, career theory has been an important guiding force for the practice of career counseling. Early theories, such as those of Ginzberg, Bordin, Roe, and Super, all helped to provide theoretical approaches to career development. Ginzberg conceptualized the process of career development as being lifelong and subject to compromise, Bordin's conceptualization of career development was informed by a psychoanalytic view of human behavior. Roe built a theory based on Maslow's hierarchy of needs and personality types. Super conceptualized career development as a developmental process and not a one-time choice. All of these early theoretical contributions still have impacts on the way in which career counselors understand human behavior today.

Arguably, the theory that has spurred the most research, and has been used the most in practice, is that of Holland. Holland articulated his comprehensive trait–factor theory in his book *Choosing a Vocation,* which was published in 1973 and revised in 1985 and 1997. Holland's theory proposed that individuals' personalities and work environments can be categorized into one of six types: Realistic, Investigative, Artistic, Social, Enterprising, or Conventional. This RIASEC theory of career development maintains that individuals will search for environments that best fit their personalities and will derive more satisfaction in finding this fit. Holland's theory has been used extensively in the United States and abroad to provide a framework for career counseling.

More recently, a number of additional theories of career development have been proposed, including Lofquist and Dawis's theory of work adjustment, Krumboltz's social learning theory of career decision making, Gottfredson's theory of circumscription and compromise in career choice, and Lent Brown and Hackett's social–cognitive theory.

5. RESEARCH IN CAREER COUNSELING

The field of career development has been one of the most actively and rigorously researched fields within counseling psychology. The field has a great deal of information available indicating that career development services, such as workshops, classes, groups, computerized career information, and guidance systems, tend to be highly effective in helping individuals to become clearer and more confident about their career goals. Considerably less information is available specifically about the efficacy of career counseling. Even less is known about what specific aspects of the career counseling process lead to effective outcomes. A recent meta-analytic study conducted by Brown and Krane analyzed 62 studies and approximately 8000 participants in attempting to determine which interventions led to the greatest change. From these data, five specific components of career counseling were shown to contribute significantly to the effect size: written exercises, individualized interpretations of information, the provision of information on the world of work, opportunities for modeling, and the building of support for career choices in individuals' social networks. The authors noted that these five interventions collectively seemed to be associated with remarkable,

nearly linear increases in career choice effect size. Thus, as the methodological tools available to researchers are becoming more sophisticated, we are learning more about what specific interventions are most helpful in the career counseling process.

See Also the Following Articles

Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Vocational Assessment in Schools ■ Vocational Interests ■ Vocational Psychology, Overview

Further Reading

Blustein, D. L., & Spengler, P. M. (1995). Personal adjustment: Career counseling and psychotherapy. In W. B. Walsh, & S. H. Osipow (Eds.), *Handbook of vocational psychology* (2nd ed., pp. 295–330). Hillsdale, NJ: Lawrence Erlbaum.

Brown, S. D., & Krane, N. E. R. (2000). Four (or five) sessions and a cloud of dust: Old assumptions and new observations about career counseling. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed., pp. 740–766). New York: John Wiley.

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, 28, 545–579.

Gysbers, N. C., Heppner, M. J., & Johnston, J. A. (2003). *Career counseling: Process, issues, and techniques*. Boston: Allyn & Bacon.

Heppner, M. J., & Heppner, P. P. (2003). Identifying process variables in career counseling: A research agenda. *Journal of Vocational Behavior*, 62, 429–452.

Holland, J. L. (1997). *Making vocational choices*. Odessa, FL: Psychological Assessment Resources.

Krumboltz, J. D. (1998). Serendipity is not serendipitous. *Journal of Counseling Psychology*, 45, 390–392.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45, 79–122.

Leong, F. L. (Ed.). (1995). *Career development and vocational behavior of racial and ethnic minorities*. Hillsdale, NJ: Lawrence Erlbaum.

Super, D. E. (1990). A life-span, life-space approach to career development. In D. Brown, L. Brooks, & Associates (Ed.), *Career choice and development* (2nd ed., pp. 197–261). San Francisco: Jossey–Bass.

Swanson, J. L. (1995). The process and outcome of career counseling. In W. B. Walsh, & S. H. Osipow (Eds.), *Handbook of vocational psychology* (2nd ed., pp. 217–259). Hillsdale, NJ: Lawrence Erlbaum.

Cheating in Academics

Gregory J. Cizek

University of North Carolina, Chapel Hill, North Carolina, USA

GLOSSARY

achievement test A procedure designed to measure an examinee's attainment of knowledge or skills.

cheating Any intentional action or behavior that: violates the established rules governing the administration of a test or the completion of an assignment; cheating gives one student an unfair advantage over other students on a test or an assignment and decreases the accuracy of the intended inferences arising from a student's performance on a test or an assignment.

crib sheet A piece of paper or other material on which inappropriate or unauthorized material is written for use during a test; it is also known as a cheat sheet.

high-stakes test A test for which serious consequences are associated with performance.

honor code A formal statement or policy concerning the rights and responsibilities of students and teachers in academic settings as well as the penalties for failing to uphold those responsibilities; the statement may be limited to matters of academic integrity but may also deal with honesty, respect, and citizenship in the academic community more broadly.

plagiarism Representing, in written or oral expression, the words, thoughts, or ideas of another person without appropriate citation or referencing, usually for the purpose of intentionally misrepresenting such words, thoughts, or ideas as one's own.

proctoring The process of monitoring test takers during an examination; a proctor (usually the instructor in a classroom) observes examinees during a test to discourage cheating and document irregularities or suspicions of cheating, but the proctor can also provide legitimate assistance to test takers if special needs, questions, and the like arise during the test.

razoring The removal of material from a printed source document for the purpose of preventing others from gaining access to the source; the term comes from the use of a razor blade to excise pages from, for example, a book or journal held in a library reserve section.

surrogate One who takes a test in place of another.

validity The degree to which the conclusions yielded by any sample of behavior (e.g., test, assignment, quiz, term paper, observation, interview) are meaningful, accurate, and useful; it is the degree to which a student's performance results in decisions about the student that are "correct" or to which inferences about the student's knowledge, skill, or ability that are "on target."

Cheating in academics refers to actions on the part of students that violate the explicit rules or commonly accepted norms for examinations or assignments. Cheating occurs at all levels, from the primary grades though testing for licensure or certification in a profession. Although most research on the topic of

cheating has focused on inappropriate behavior on the part of those who take tests, cheating in academics can also include cheating by those who give tests. Reasonably accurate methods exist for detecting some types of cheating. Many strategies have also been proven to be effective for organizing instruction and assessment so that cheating can be prevented.

1. INTRODUCTION

Sound assessment practice involves the development and administration of high-quality instruments and procedures to measure student learning. Assessment instruments can include graded tests and assignments. Information that results from these instruments can be helpful to students themselves or to others who use the information for educational decision making such as teachers, administrators, and policymakers. Increasingly in education contexts, consequences are also associated with performance on assessments. When the associated consequences are relatively serious, the assessment is said to involve high stakes.

The stakes associated with an assessment can be serious or mild, and they can be positive or negative. For example, awarding scholarship money to a student who performed at an exceptional level on a college admissions test would be considered a serious positive consequence. Assigning an elementary student to a remedial reading program due to a low test score could be either a positive or a negative consequence, depending on whether or not the remedial program was effective and other factors. Withholding a diploma because a student failed a graduation examination would be a fairly serious negative consequence. As the importance of decision making increases, and as the stakes associated with assessment rise, the problem of cheating becomes more pronounced.

Studies of cheating by students indicate that the incidence of the behavior increases as students pass through the grade levels, with the greatest frequency at the high school and college levels. Research evidence indicates that approximately 90% of students admit to having cheated on a test or an assignment, and it is estimated that 3 to 5% of students cheat on a given test or assignment.

2. DEFINITIONS AND EXAMPLES

This article focuses on the kinds of cheating that occur in classroom situations. Cheating can occur on tests,

for example, when one test taker copies test answers from another test taker. Cheating may also occur on assignments, such as on term papers or projects, when a student inappropriately uses material taken from another source for the assignment.

Regardless of the situation, a general definition of cheating can be developed. Cheating can be defined as any intentional action or behavior that violates the established rules governing the administration of a test or the completion of an assignment, gives one student an unfair advantage over other students on a test or an assignment, or decreases the accuracy of the intended inferences arising from a student's performance on a test or an assignment.

Varieties of cheating by test takers can range from simple copying of another student's answers to more elaborate schemes. An examinee may use unauthorized materials, such as a cheat sheet, or may take advantage of the testing situation by requesting testing accommodations that are not necessary. Cheating on assignments may involve impermissible books, notes, or persons when the use of such aids is contrary to guidelines provided for an out-of-class assignment. Academic cheating also may include plagiarism, in which the original words, thoughts, or ideas of another person are used without appropriate citation or referencing (usually for the purpose of intentionally misrepresenting such words, thoughts, or ideas as one's own). Many Internet sources exist from which students can, either at no charge or for a fee, copy or download complete term papers or smaller portions of existing works. Finally, other forms of cheating exist in testing situations (e.g., when a person does not actually take a test himself or herself but rather relies on another person, called a confederate or surrogate, to do so) or on assignments when cheating may take the form of razoring (which involves the removal of material from a printed source document for the purpose of preventing others from gaining access to the source). More extensive lists of cheating methods were provided by Cizek in 1999 and 2003.

3. CHEATING AND VALIDITY

Validity is the single greatest concern in any assessment situation. The concept refers to the accuracy of the interpretations about examinees based on their performance on a test or an assignment. In technical terms, validity is the degree to which evidence supports the inferences made about a person's knowledge,

skill, or ability based on his or her observed performance. By definition, inferences are based on a less than ideal amount of information such as a sample of a student's knowledge or skill obtained via a test or term paper. Because it is often too costly or impractical to gather more information, inferences must be based on samples of behavior. Consequently, it is necessary to consider the accuracy of inferences based on the available evidence (e.g., test performance), that is, to consider validity. Any factor that attenuates the ability to make accurate inferences from the sample of performance threatens validity and jeopardizes the meaningfulness of conclusions based on the sample at hand. When cheating occurs, inaccurate inferences result.

4. PROFESSIONAL GUIDELINES ON CHEATING

Many educational institutions have developed honor codes to formalize expectations about academic integrity; to clearly define the types of actions that constitute cheating on tests, assignments, or other academic work; and to outline penalties for cheating. Test publishers usually produce carefully scripted directions for administering their tests and provide clear guidelines for what kind of behavior is permissible and what kind is not. Acceptable behavior and unacceptable behavior are also sometimes codified in states' administrative regulations or statutes. Assessment specialists, through their professional associations, have developed professional guidelines to inform test takers and test administrators regarding inappropriate practices. The most authoritative of these guidelines are found in the *Standards for Educational and Psychological Testing,* a document that focuses on testing. Among their provisions, the standards indicate that those involved in testing programs should

- protect the security of tests (Standard 11.7);
- inform examinees that it is inappropriate for them to have someone else take the test for them, disclose secure test materials, or engage in any other form of cheating (Standard 8.7);
- ensure that individuals who administer and score tests are proficient in administration procedures and understand the importance of adhering to directions provided by the test developer (Standard 13.10);
- ensure that test preparation activities and materials provided to students will not adversely affect the

validity of test score inferences (Standard 13.11); and
- maintain the integrity of test results by eliminating practices designed to raise test scores without improving students' real knowledge, skills, or abilities in the area tested (Standard 15.9).

5. DETECTING CHEATING

Attempts to detect cheating can be performed by observational/judgmental or statistical methods. Observational/Judgmental methods have the potential to detect cheating on academic assignments and tests. Statistical methods have been developed to detect a particular type of cheating (copying) on tests.

Observational/Judgmental methods rely more heavily on subjective human perceptions. For example, a student might enlist the aid of a surrogate to take an examination in his or her place. Human judgment is involved in detecting this form of cheating; exam proctors must be alert for test takers who do not appear familiar to them, must scrutinize examinees' identifications prior to the test, and so on. Judgment is also involved when handwriting samples from the student are compared with those of the surrogate to make a determination of whose handwriting appears on the test materials or when a student obtains an unusually large score gain from one test administration to another.

Judgment is also involved in the detecting of plagiarism on written assignments such as term papers. A number of techniques can be used to ascertain whether a student's written work is original or may have been plagiarized. Table I lists some common observational/judgmental methods for detecting this form of cheating. In addition to observational/judgmental methods, statistical methods and computer technology can be invoked to aid in the detection of cheating.

Basic search engines (e.g., Google, Yahoo) can be used to locate Internet sources from which plagiarized material may have been drawn. This method relies on an instructor entering a unique word (or, better yet, a unique phrase or string of words) from a student's paper to identify Internet sources containing the same string. Commercial (i.e., fee-based) Web sites devoted to the detection of plagiarism also exist. Most such sites require student work to be submitted electronically. Then, students' papers are compared with those in databases of student work maintained by the host of

TABLE I
Observational Indications of Potential Plagiarism

1. The writing style, language, vocabulary, tone, grammar, and/or other features of a student's writing are not commensurate with what the student usually produces.
2. There are awkward changes in verb tense, pronouns, structure, and/or organization of the paper. The writing is choppy. Sections or sentences in the paper seem out of place or do not relate to the overall topic of the paper.
3. Strange text, such as a date in the footer of a paper or a URL (Web site address) in a header, appears at the top or bottom of printed pages.
4. There is unusual or inconsistent formatting such as page layout, line spacing, margins, and/or pagination. The paper contains changes in font, pitch, color, and the like, suggesting that the paper may have been downloaded from the Internet or cut-and-pasted from another source.
5. The paper contains references to graphs, charts, accompanying material, citations, chapters, footnotes, and/or additional text that are not provided with the paper.
6. The paper includes allusions to ideas, persons, settings, and/or other elements with which the student is not likely familiar.
7. The paper contains citations or references to sources or materials to which the student is not likely to have access.
8. Citations or references are missing, inaccurate, or incomplete.
9. The student has difficulty summarizing his or her paper or in responding to simple questions about what he or she wrote.

Source. Adapted from Bates and Fain (2002, p. 1) and Cizek (2003a, chap. 4).

the site. An overall probability is produced for each student's work, indicating the overall degree of similarity between the student's paper and all of the other sources. Software for local use that essentially performs a similar function of searching for common phrasings can also be purchased.

Statistical methods are also available to gauge the probability of copying on multiple-choice tests. Information on these methods, including details regarding how to compute statistical indexes of cheating, was provided by Cizek in 1999 and by Sotaridona in 2003.

6. PREVENTING CHEATING

Many strategies exist for deterring cheating in academics. Many of these strategies are unique to the particular type of cheating to be deterred (i.e., whether on a test or on a written assignment). Other strategies aim to foster higher levels of academic integrity in general. Table II provides a list of specific ways in which to prevent cheating on tests and written assignments.

In addition, at the college and university levels, many institutions have found that inappropriate behavior can be prevented by developing and disseminating explicit expectations about academic integrity. Such expectations, formalized in what are referred to as honor codes, help to promote a campus culture and community characterized by the valuing of ethical conduct. An excerpt of one such honor code is reproduced in Table III. The excerpt shown is limited to the portion of the code related to cheating, and a source reference is provided for accessing the full code.

TABLE II
Strategies for Preventing Cheating

1. Limit the amount of tests and assignments that "count" toward student grades. Foster an assessment climate in which students apprehend the value of accurate information on their strengths and weaknesses. Avoid using grades punitively.
2. Introduce students to the requirements of academic integrity explicitly and often. Provide concrete illustrations of inappropriate actions. Teach appropriate test preparation and citation practices.
3. Avoid assessment formats that make cheating easier. For example, reduce reliance on multiple-choice items in favor or constructed-response formats. Use alternate forms of tests in which test item positions are scrambled. Avoid "take home" tests. Avoid repeating term paper topics or test questions from year to year.
4. Keep test materials secure in advance of testing. Follow appropriate steps during testing such as requiring adequate spacing between test takers. Remain vigilant while proctoring.
5. Be fair and open about assessment. Provide students with adequate notice and description of examinations and adequate time to complete written assignments.
6. Inform students that statistical methods for detecting answer copying and technological aids for detecting plagiarism will be used.
7. Institute an honor code. Develop and disseminate policies on cheating. Implement penalties for academic dishonesty.

A. General Responsibilities

It shall be the responsibility of every student at the University of North Carolina at Chapel Hill to

1. Obey and support the enforcement of the Honor Code;
2. refrain from lying, cheating, or stealing;
3. conduct themselves so as not to impair significantly the welfare or the educational opportunities of others in the university community; and
4. Refrain from conduct that impairs or may impair the capacity of university and associated personnel to perform their duties, manage resources, protect the safety and welfare of members of the university community, and maintain the integrity of the university.

Offenses proscribed by this section include, but shall not be limited to, those set out in Sections II.B and II.C.

B. Academic Dishonesty

It shall be the responsibility of every student enrolled at the University of North Carolina to support the principles of academic integrity and to refrain from all forms of academic dishonesty, including but not limited to the following:

1. Plagiarism in the form of deliberate or reckless representation of another's words, thoughts, or ideas as one's own without attribution in connection with submission of academic work, whether graded or otherwise
2. Falsification, fabrication, or misrepresentation of data, other information, or citations in connection with an academic assignment, whether graded or otherwise
3. Unauthorized assistance or unauthorized collaboration in connection with academic work, whether graded or otherwise
4. Cheating on examinations or other academic assignments, whether graded or otherwise, including but not limited to the following: (a) using unauthorized materials and methods (notes, books, electronic information, telephonic or other forms of electronic communication, or other sources or methods); (b) violating or subverting requirements governing administration of examinations or other academic assignments; (c) compromising the security of examinations or academic assignments; (d) representing another's work as one's own; or (e) engaging in other actions that compromise the integrity of the grading or evaluation process
5. Deliberately furnishing false information to members of the university community in connection with their efforts to prevent, investigate, or enforce university requirements regarding academic dishonesty
6. Forging, falsifying, or misusing university documents, records, identification cards, computers, or other resources so as to violate requirements regarding academic dishonesty
7. Violating other university policies that are designed to ensure that academic work conforms to requirements relating to academic integrity
8. Assisting or aiding another to engage in acts of academic dishonesty prohibited by Section II.B.

Source. University of North Carolina at Chapel Hill (2003).

See Also the Following Articles

Achievement Tests ■ Educational Achievement and Culture ■ Educational and Child Assessment ■ Psychological Assessment, Standards and Guidelines for

Further Reading

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Bates, P., & Fain, M. (2002). *Detecting plagiarized papers.* www.coastal.edu.

Cizek, G. J. (1999). *Cheating on tests: How to do it, detect it, and prevent it.* Mahwah, NJ: Lawrence Erlbaum.

Cizek, G. J. (2003a). *Detecting and preventing classroom cheating: Promoting integrity in schools.* Thousand Oaks, CA: Corwin.

Cizek, G. J. (2003b). Educational testing integrity. In J. E. Wall, & G. R. Walz (Eds.), *Measuring up: Assessment issues for teachers, counselors, and administrators* (pp. 363–386). Greensboro, NC: ERIC Counseling and Student Services Clearinghouse.

Sotaridona, L. S. (2003). *Statistical methods for the detection of answer copying on achievement tests.* Enschede, Netherlands: Twente University Press.

University of North Carolina at Chapel Hill. (2003). *Instrument of student judicial governance.* Chapel Hill, NC: Author. (Available: http://instrument.unc.edu)

Cheating in Sport

Glyn C. Roberts, Yngvar Ommundsen, Pierre-Nicolas Lemyre and Blake W. Miller

Norwegian University of Sport Science, Oslo, Norway

Recent research reveals that understanding the achievement goal of a person may help to explain cheating in sport. Achievement goals have dispositional elements as well as situational determinants, and both have been associated with cheating and moral action in sport. The more ego involved the person, the more the person cheats. Reducing cheating may be achieved through coach education and by deemphasizing normative comparisons.

GLOSSARY

achievement goals Within an achievement context, the goals that give meaning to achievement striving.

cheating A quest to provide an unfair advantage over the opponent; cheating may be "unsportspersonlike" aggressive behavior or inappropriate moral action, but always with the intent to gain an advantage.

dispositional goals Achievement goals that are within a person, that is, an individual difference variable.

ego involvement When the goal of a person is to demonstrate superiority to peers.

moral atmosphere The collective norms of the team about whether or not to condone cheating.

motivational climate The perceived criteria of success and failure that a person perceives are manifest within an achievement context.

task involvement When the goal of a person is to achieve mastery or learning.

1. INTRODUCTION

Cheating in sports is endemic. Each week brings new revelations about cheating, from sprinters who take banned substances (e.g., Dwain Chambers, Carl Lewis, Kelly White) to endurance athletes who take performance-enhancing drugs (e.g., Roberta Jacobs, Richard Virenque, Pamela Chepchumba). But cheating is not limited to individual athletes; prestigious sport organizations have also been caught trying to bend the rules for their own advantage. Some of the most blatant examples of recent cases include the following. Six athletes from the Finnish Ski Federation were discovered to have taken performance-enhancing drugs with the blessing of the medical staff and were disqualified at the Lahti World Cup in 2001. The South Africa, Pakistan, and India cricket teams were found to have fixed matches in 2000. The Welsh Rugby Association recruited illegal players to play for Wales in 2000 by "discovering" fictitious Welsh grandparents for the players. Cyclists on sponsored teams

(e.g., the Festina cycling team) in the Tour de France were caught using banned substances in 2001 and 2002. Cheating has even polluted children's sports, with the most recent visible case being that of Little League pitcher Danny Almonte, who was 14 years of age when he was the winning pitcher for the Little League World Series (restricted to children 9–12 years of age) in 2001. Danny's father has been indicted for the fraud.

At the time of this writing, a huge cheating scandal involving track and field athletes was breaking in the United States. On October 16, 2003, the U.S. Anti-Doping Agency (USADA) released a statement describing the discovery of a new "designer steroid" that had been deliberately manufactured by BALCO laboratories in the United States to avoid detection by current testing procedures. The new steroid (tetrahydrogestrinone or THG) was discovered only after a prominent coach sent a syringe containing the substance to the USADA and named several prominent athletes who were using the steroid. The anti-doping laboratory at the University of California, Los Angeles, identified and developed a test for the substance, and the USADA retrospectively tested the samples from the U.S. Outdoor Track and Field Championships in 2003. Several positive sample results have emerged, and it has been decided that all of the samples from the 2003 World Championships in Paris are to be tested. Terry Madden, the USADA chief executive, stated that this is a case of intentional doping of the "worst sort" and that it involved a conspiracy involving chemists, coaches, and athletes in a deliberate attempt to defraud fellow competitors and the world public. The testing of both "A" samples and, if positive, subsequent "B" samples was ongoing at the time of this writing. In addition, Victor Conte, the founder of the BALCO laboratory that is at the heart of the scandal, was indicted to appear before a grand jury investigating the illegal distribution of a controlled substance. The scandal could involve many athletes and will tarnish many reputations and the sport. Among the athletes named thus far are U.S. shot put champion Kevin Roth, hammer thrower John McEwan, women's hammer throw champion Melissa Price, and middle distance runner Regina Jacobs, who all tested positive at the 2003 U.S. championships. The European sprint champion, the United Kingdom's Dwain Chambers, has also tested positive. Among other "clients" of BALCO laboratories who have been implicated are sprinters Marion Jones and Tim Montgomery, baseball star Barry Bonds, and the NFL's Bill Romanowski.

Not all cheating is as obvious. Each sport has its own manifestation of illegal or inappropriate activity. As an example, professional soccer has what is termed the "professional foul" where players will sometimes pull an opponent's shirt to prevent or hinder progress with the ball. This is now so common that referees do not call it unless it is very blatant. Football players also "dive" when challenged for the ball, trying to draw a penalty. Baseball players cork their bats. The world's number one golfer, Tiger Woods, has accused many of his fellow professionals of using illegal drivers that give a "trampoline effect" to the ball. Cricket players tamper with the ball to make it "swing" more through the air. And so on,

Why do people cheat in sport? Many explanations have been given, and they are mainly economic (the rewards of being a successful elite athlete are huge), sociological (a breakdown in the moral fabric of society), and psychological (the focus of the current article). However, it is the act of competing that has been indicated the most. It is argued that competition, whether on the sport field, in the classroom, or in the boardroom, is the culprit. When people compete, especially in important contexts, people will cheat to achieve success. When winning is everything, they will do anything to win. But so many people who compete do not cheat, so it might not be competition per se that is the problem. An interesting question may then be asked: What are the determinants of cheating, and how are these manifested in sport contexts? The current authors' research over the past few years has investigated this issue and, like so many other researchers investigating cheating in sport, has looked at children's sport. What are some of the determinants of children engaging in inappropriate behavior in sport? First, one must look at the context of sport, and what it means for children, and then define what is meant by "cheating" before discussing the manifestation of cheating in sport.

2. THE CONTEXT OF SPORT

Performance in sport is clearly important for adult athletes; the rewards of competence can be significant. In addition, public interest in sport is high and is reflected in the column inches devoted to sport in the popular newspapers as well as in television airtime. It was estimated that one in four persons worldwide watched the final World Cup soccer match between Brazil and Germany on television in 2002. Thus, the entertainment value of sport is increasing for adults. The context is also important to children for different reasons.

Performing in sport contexts is assumed to be important in the socialization process of children toward

the development of appropriate moral behavior. In play, games, and sport, children are brought into contact with social order and the values inherent in society and are provided a context within which desirable social behavior is developed. The psychosocial and moral development of young participants is fostered when peer status, peer acceptance, and self-worth can be established and developed, and the adoption of various perspectives is enhanced. Sport is also assumed to provide a vehicle for learning to cooperate with teammates, negotiate and offer solutions to moral conflicts, develop self-control, display courage, and learn virtues such as fairness, team loyalty, persistence, and teamwork. Despite popular beliefs that sport builds character, this notion has been questioned. Research has shown that competition may promote antisocial behavior and reduce prosocial behavior.

The context of sports is becoming an increasingly important one for modern-day children. With the demise of children's game-playing culture, children are more and more likely to be involved with adult organized sport competition, even as young as 4 years of age (e.g., motocross in Belgium). In addition, research has demonstrated that the domain of competitive sport is a particularly important context for psychosocial development in that peer status, peer acceptance, and self-worth are established and developed. These social attributes are based on many factors, but one way in which a child can gain peer acceptance and status is to demonstrate competence in an activity valued by other children. One area of competence that is highly valued by children is sporting ability. In fact, because of the modern-day visibility of sport, being a good sport player appears to be a strong social asset for a child, especially in the case of boys. Thus, the context is an important one and is one where cheating to gain advantage or gain acceptance with one's peers can take place.

3. WHAT IS CHEATING?

In the area of sport moral action, especially with children, many variables have been investigated. Not all of them may be defined as cheating per se. For example, moral atmosphere has been included and refers to the cultural norms developed within a team about whether cheating is condoned or not. However, all of the variables may be defined as inappropriate behavior, at the very least, and range from poor "sportspersonship" to outright aggression so as to achieve a competitive advantage. Because cheating is a difficult concept to

pinpoint and define universally, it may be helpful to view cheating as a product or combination of several more or less moral concepts that are referred to often in the sport psychology literature. For example, cheating in sport has been viewed as a quest to provide an unfair advantage over the opponent. According to this definition, a number of concepts fall into the category of cheating in sport, for example, unsportspersonlike aggressive behavior, inappropriate moral functioning, and a team moral atmosphere where cheating is condoned.

Of all the virtues that sport supposedly fosters, sportspersonship is perhaps the most frequently cited. The virtue of sportspersonship is oriented toward maximizing the enjoyable experience of all participants. Although most people believe that they know what sportspersonship is, the development and understanding of the concept has suffered from the lack of a precise definition and an overreliance on broad theoretical approaches. In essence, a sport participant manifests sportspersonship when he or she tries to play well and strive for victory, avoids taking an unfair advantage over the opponent, and reacts graciously following either victory or defeat.

In an effort to generate a much-needed conceptual base to promote research, recent work from Vallerand and colleagues has contributed to a better understanding of the sportspersonship concept. Vallerand has adopted a social psychological view of sportspersonship that separates the latter from aggression and assumes a multidimensional definition consisting of five clear and practical dimensions: full commitment toward sport participation, respect for social conventions, respect and concern for the rules and officials, true respect and concern for the opponent, and negative approach toward sportspersonship. This definition and instrument has been used in sport to investigate sportspersonship. Research has used the scale; however, the negative dimension has never really worked out and is frequently not included. Therefore, when children respect the social conventions (e.g., shaking hands after a game), respect the rules and officials (e.g., not violating the rules or arguing with the referee), and respect the opponent (e.g., helping an opposing player up from the ground in football), they are considered to be high in sportspersonship.

Sportspersonship is but one dimension of social moral functioning in sport, and one model used to more fully investigate the issue is a model suggested by Rest in 1984. Rest proposed a four-component interactive model of sociomoral action that seems well suited when examining sociomoral aspects in

sport. The first component of the model deals with interpreting the sport situation by recognizing possible courses of moral action. The second component encompasses forming a moral judgment involving judgments about both the social legitimacy and moral legitimacy of inappropriate sport behavior. The third component involves deciding what one intends to do as a solution to the dilemma, and the fourth involves executing and implementing one's intended behavior. Behavior might include incidences of nonmoral actions (e.g., aggression) as well as sportspersonship behavior.

The process of making a moral decision may be influenced by many factors, including motivational factors, whereas actual behavior may be affected by distraction, fatigue, and factors that physically prevent someone from carrying out a plan of action. The interactive nature of the four processes means that factors proposed to act primarily on one process also influence the others indirectly. As an example, when a child is in a competitive context, many kinds of behavior are possible, including whether to cheat or not. A situation may develop where one could cheat to stop a player from scoring, but is it appropriate to do so? If doing so is judged as appropriate, does one intend to cheat to stop the player from scoring? Finally, does one actually cheat to stop the player from scoring? In this way, players go through the components of Rest's model.

In 1985, Shields and Bredemeier argued that a major factor influencing the construction of a moral judgment, and consequently moral behavior, in sport is the moral atmosphere of the team. To capture the perceived judgments by significant others as either approving or disapproving of one's moral actions, research has often included how the players perceived the moral atmosphere in terms of social moral team norms. Sport teams, like all groups, develop a moral atmosphere composed of collective norms that help to shape the moral actions of each group member. An example again is the professional foul in soccer. Sociomoral atmosphere, or perceived sociomoral team norms, has been measured by means of a questionnaire.

Another variable has been the perceived legitimacy of intentionally injurious acts. Participants are asked to imagine themselves competing in an important competitive context. Various intentionally injurious acts are described, and participants are asked to indicate to what degree they agree or disagree with the legitimacy of the actions described.

This is how moral functioning and cheating have been defined in research in sport, especially with children. But why do people cheat? One avenue for exploring this issue has been to look at the motivation to cheat and has used concepts from motivation theory and research. The basic arguments behind this line of research are considered next.

4. THE MOTIVATION TO CHEAT: ACHIEVEMENT GOALS AS A DETERMINANT OF MORAL ACTION

The most promising avenue to investigate moral functioning and action in sport has been to use achievement goal theory. This framework assumes that achievement goals govern achievement beliefs and guide subsequent decision making and behavior in achievement contexts. It is argued that to understand the motivation of individuals, the function and meaning of the achievement behavior to the individual must be taken into account and the goal of action must be understood. It is clear that there may be multiple goals of action rather than just one. Thus, variation of behavior might not be the manifestation of high or low motivation per se; instead, it might be the expression of different perceptions of appropriate goals. An individual's investment of personal resources, such as effort, talent, and time in an activity as well as moral action, may be dependent on the achievement goal of the individual in that activity.

The goal of action in achievement goal theory is assumed to be the demonstration of competence. In 1989, Nicholls argued that two conceptions of ability manifest themselves in achievement contexts: an undifferentiated concept of ability (where ability and effort are perceived as the same concept by the individual or he or she chooses not to differentiate) and a differentiated concept of ability (where ability and effort are seen as independent concepts). Nicholls identified achievement behavior using the undifferentiated conception of ability as task involvement and identified achievement behavior using the differentiated conception of ability as ego involvement. The two conceptions of ability have different criteria by which individuals measure success. The goals of action are to meet those criteria by which success is assessed. When task involved, the goal of action is to develop mastery or improvement and the demonstration of ability is self-referenced. Success is realized when mastery or improvement has been attained. The goal of action for an ego-involved individual, on the other hand, is to demonstrate normative ability so as to outperform others. Success is realized when the performance of others is exceeded, especially when little effort is expended.

Whether one is engaged in a state of ego or task involvement is dependent on the dispositional orientation of the individual as well as situational factors. Consider the dispositional aspect first. It is assumed that individuals are predisposed to act in an ego- or task-involved manner. These predispositions are called achievement goal orientations. An individual who is task oriented is assumed to become task involved, or chooses to be task involved, so as to assess demonstrated competence in the achievement task. The individual evaluates personal performance to determine whether effort is expended and mastery is achieved; thus, the demonstration of ability is self-referenced and success is realized when mastery or improvement is demonstrated. In contrast, an individual who is ego oriented is assumed to become ego involved in the activity. The individual evaluates personal performance with reference to the performance of others; thus, the demonstration of ability is other-referenced and success is realized when the performance of others is exceeded, especially when little effort is expended.

Achievement goal theory holds that the state of motivational goal involvement that the individual adopts in a given achievement context is a function of both motivational dispositions and situational factors. An individual enters an achievement setting with the disposition tendency to be task and/or ego oriented (goal orientation), but the motivational dynamics of the context will also have a profound influence on the adopted goal of action, especially for children. If the sport context is characterized by a value placed on interpersonal competition and social comparison, the coach emphasizing winning and achieving outcomes, and a public recognition of the demonstration of ability, a performance climate prevails. This reinforces an individual's likelihood of being ego involved in that context. If, on the other hand, the context is characterized by learning and mastery of skills, trying hard to do one's best, and the coach using private evaluation of demonstrated ability, a mastery climate prevails. An individual is more likely to be task involved in that context. Therefore, being task or ego involved is the product of an interaction of personal dispositions and the perceived motivational climate. However, in the research literature, investigators typically look at one aspect or the other are rarely look at them in combination.

Shields and Bredemeier argued that situational influences may have a great effect on an athlete's moral action. Competitive ego-involving structures may focus the individual's attention on the self and, in the case of team sports, on those comprising the in-group as well. This may reduce the player's sensitivity to the welfare of opposing players. Extensive involvement in competitive contexts may reduce the person's ability to show empathy, thereby reducing consideration for the needs of others faced with a moral dilemma. Indeed, several studies have shown that participation in competition is associated both with reduced sportspersonship and prosocial behavior and with increased antisocial behavior, hostility, and aggressiveness. However, it is argued here that it might not be competition in and of itself that induces sociomoral dysfunction on the individual and group levels. Rather, it might be the perceived motivational climate that may shape an athlete's moral functioning. Indeed, moral development theorists, such as Rest, agree that moral behavior is intentional motivated behavior. Thus, to predict sociomoral perceptions and actions, one must consider the motivational characteristics of the situation. First, dispositional influences on cheating are examined.

5. DISPOSITIONAL ACHIEVEMENT GOALS

When looking at the various important personal factors influencing people's behavior in sport, moral reasoning ability and achievement goal orientations have emerged as significant variables. An early research study found that moral reasoning in the context of sport is much more egocentric than moral reasoning in most situations in everyday life. Athletes, as compared with nonathletes, seem to change their basis for moral reasoning from non-sport- to sport-related activities. This type of reasoning seems to focus on self-interest and personal gain.

Research has shown that when athletes are task or ego oriented, they differ in various aspects of moral behavior as well as in views regarding what represents acceptable behavior in sport. For example, athletes who are primarily concerned with outperforming others (ego oriented) have been found to display less mature moral reasoning, tend to be low in sportspersonlike behavior, approve of cheating to win, and perceive intentionally injurious and aggressive acts in sport as legitimate. In contrast, among task-oriented athletes who tend to use self-referenced criteria to judge competence and feel successful when they have achieved mastery or improvement of the task, greater approval of sportspersonship, disagreement with cheating to win, and less acceptance of intentionally injurious acts have been observed.

Research has found that players who reported higher temptation to play unfairly and greater approval of

behavior designed to obtain an unfair advantage were more likely to be ego oriented. These players also perceived their coaches as being more ego involving than task involving. In addition, the athletes believed that teammates would play more unfairly and would have a higher rate of approval of behavior aimed at obtaining an unfair edge over the opponent. In a study investigating the influence of moral atmosphere (i.e., the collective team norms approving of cheating) on the moral reasoning of young female soccer players, players who were in a culture where aggression and unfair play were tolerated were more ego oriented. Furthermore, it has been found that ego-oriented players had lower levels of moral functioning, had greater approval of unsportspersonlike behavior, and judged that intentionally injurious acts are legitimate. Ego orientation has been related to unacceptable achievement strategies such as the use of aggression. Findings indicate that players high in ego orientation display more instrumental aggression than do participants with low ego orientation, suggesting that participants high in ego orientation may adopt a "win at all costs" attitude.

However, ego-involved individuals can also adopt adaptive achievement strategies. It is believed that when individuals are both ego involved and high in perceived ability, so long as the perception of high ability lasts, these individuals will seek challenging tasks, will revel in demonstrating their ability, and generally will not cheat. But when these individuals experience a change in their perception of ability, they are more likely to cheat so as to continue winning. Research has confirmed this in that high ego-, low task-oriented participants who are low in perceived ability are more likely to endorse cheating in sport than are high ego-, low task-oriented participants who are high in perceived ability. High task-oriented individuals are less likely to approve of cheating in sport than are any other sport participants, regardless of perceived ability.

An athlete high in ego orientation is assumed to believe that winning is the single most important thing when competing. When doubt about one's own ability exists, this may lead the young athlete to believe that it is harder to contribute to his or her team's success by using skills alone. The athlete can then be tempted to cheat so as to help the team win and gain peer acceptance. In some of the current authors' research over the years with young players in soccer, they have found that an ego achievement goal orientation is associated with poor sportspersonship and cheating when perception of ability is low. These findings suggest that young athletes are likely to cheat when they do not believe that they can succeed by simply playing fair. How far one is willing to go to win is often seen as a positive attribute in the world of competitive sport. If one player is willing to get caught or be reprimanded by the officials while trying to intimidate or undermine the opponent, this athlete is likely to generate a positive response from his or her peers, parents, and coaches when playing on a team where winning is everything. Team loyalty can be demonstrated by going against the rules, particularly if the athlete does not get caught. The current authors have found that athletes with a high ego orientation show low respect for rules and officials as well as low respect for opponents.

When winning is everything and perception of ability is low, cheating and going against the rules may be judged to be acceptable so as to gain peer acceptance. This is true for both boys and girls. But boys cheat more. To illustrate this, two figures are included showing how boys and girls differ in degree if not in kind. These figures present data from the authors' own ongoing research involving young soccer players in the Norway Cup, the largest international competition for children in the world. The data show the impact of being high in ego orientation. In this particular example, the authors, using the Rest model to inform their questions, asked the players some questions about their experience with playing soccer. They wanted to see the relationships of the motivational variables on moral reasoning and moral action variables. Figure 1 shows a canonical correlation chart of the responses of high-ego girls (ages 15–16 years). To be meaningful, the correlations should be greater than .30. As can be seen in the figure, when girls are ego involved, they judge inappropriate behavior as being appropriate, and the most important reason was negative; therefore, they were low in moral reasoning, they intended to cheat, and they judged injurious acts as being appropriate to win.

Not attaining significance, these girls had not cheated over the previous five games and considered the team atmosphere and the coaching climate as being relatively neutral. However, they were quite prepared to injure an opponent in the quest to win. Compare this with Fig. 2, which shows a canonical correlation chart of the responses of high-ego boys. These boys judged inappropriate behavior as being appropriate, and the most important reason was clearly negative; therefore, they were low in moral reasoning, and all of the other variables are clearly positive.

This means the high-ego boys admitted that they had cheated over the previous five games, thought that the team atmosphere supported cheating to win, considered the coach as supporting cheating, and were quite
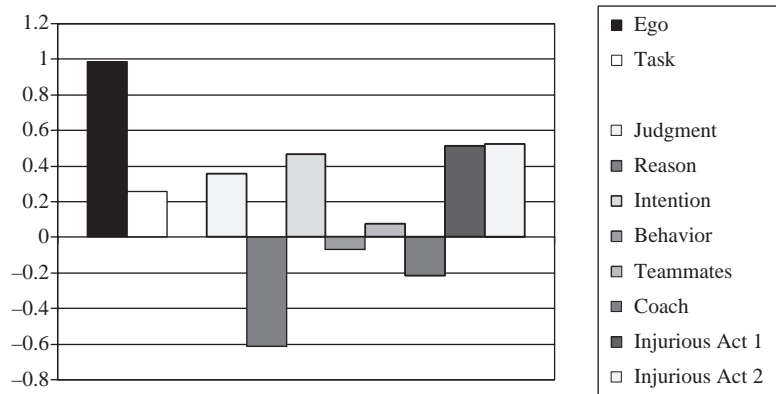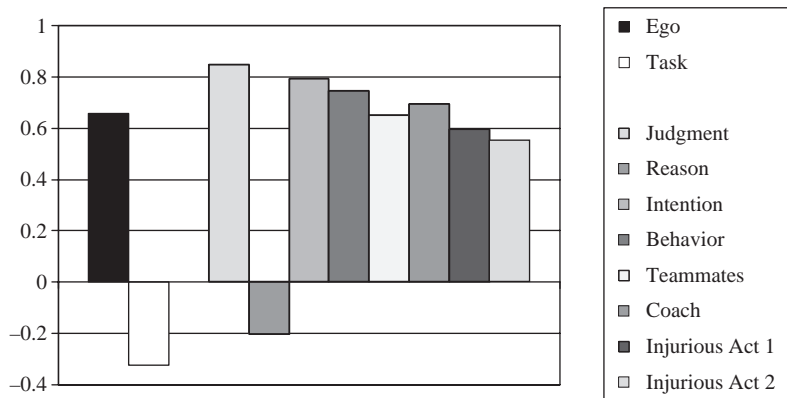
FIGURE 1   Girls: High ego, low task.

FIGURE 2   Boys: High ego, low task.

prepared to injure opponents in their desire to win. These data show clearly that high-ego players have more tolerance toward and support cheating to win, even to the extent of supporting injurious acts. Although the boys and girls thought and behaved in the same general way, boys are more extreme and do cheat more.

Clearly, being ego or task involved has implications for moral action in sport. The influence that the motivational climate has on moral thinking and action is considered next.

6. PERCEIVED MOTIVATIONAL CLIMATE AND DETERMINANTS OF CHEATING

Similar to the task- and ego-involving criteria to which an individual may be disposed, the situational aspects of the competitive environment are also important. Although a number of significant others (e.g., teachers, parents, peers) can and do influence, to some extent, the type of motivational climate that is perceived, the coach is perhaps the most influential facilitator of task- or ego-involving criteria for the athlete. The bottom line here is that the criteria of success and failure that the coach brings into the situation are the criteria that the athlete adopts. The younger the athlete, the more likely it is that he or she will pay attention to the criteria of the coach. Indeed, there is evidence indicating that the criteria of what constitutes success and failure for the coach "washes out" the personal criteria of the young player. As such, the achievement goal that the coach imposes on the context becomes an important determinant of the behavior of the athlete.

In discussing the motivational climate, the terms "performance climate" (an ego-involving coach) and "mastery climate" (a task-involving coach) are used. If a performance climate prevails, the coach and the

athletes may come to view competition as a process of striving against others. The players may perceive pressure from the coach to perform well or be punished as well as pressure to outplay opponents and win so as to receive recognition and attention. Players may resort to cheating, violating rules, and behaving aggressively as a means of coping. In competition, a performance climate may generate a strong interteam rivalry. A hostile atmosphere toward opposing players may result, leading to the development of team norms or shared social moral perceptions among team members reflecting a derogatory and depersonalized picture of opposing players as mere obstacles to be overcome in the quest for victory by their own team.

The investigation of perceived motivational climate in relation to cheating is a relatively new research area among sport social scientists. The findings thus far indicate that perceived mastery and performance motivational climates created by coaches do affect cheating in systematic ways.

Players perceiving a strong mastery motivational climate reported that sportspersonlike behavior is important in competitive soccer. Players reported that it was important to value and respect the rules of the game and the officials who represented the rule structure. In addition, players noted the importance of respecting the social conventions found within the soccer environment when they perceived the coach as emphasizing mastery motivational criteria. In contrast, it has been shown that players who perceive a high performance climate indicate lower sportspersonship than do players who perceive a high mastery climate.

In general, the majority of research indicates that a perceived performance motivational climate is associated with cheating, whereas a perceived mastery motivational climate is associated with more positive moral actions. This is true for college-level athletes as well as for younger athletes. In brief, perceptions of performance climate criteria were related to a lower level of moral understanding within the team, resulting in lower moral functioning illustrated by means of low moral judgment, a high intention to behave nonmorally, and low self-reported moral behavior.

Among high-level collegiate male and female basketball players, it has been found that players perceiving a performance motivational climate perceived a low moral atmosphere within their team that was related to lower moral functioning. Various perceptions of the motivational climate have been shown to be related to differing levels of moral functioning and perceptions of team moral atmosphere. Evidence illustrated in Fig. 3 shows that young players who perceive a strong mastery motivational climate indicated positive sportspersonship (i.e., had respect for opponents and officials and also respected social conventions), had positive moral reasoning, had a low intention to cheat, seldom engaged in cheating behavior, and considered the team norms and the atmosphere created by the coach to be anti-cheating.

In contrast, Fig. 4 shows a canonical correlation chart of the responses of girls in a performance-oriented climate. When girls are ego involved, they judged inappropriate behavior as being appropriate, and the most important reason was negative; therefore, they were low in moral reasoning, intended to cheat, had cheated during the previous five games, considered that the coach and teammates approved of cheating, and considered injurious acts appropriate to win. Again, the figure illustrates that even girls (and young female youth), when ego involved through a performance climate created by the coach, will increase their cheating.
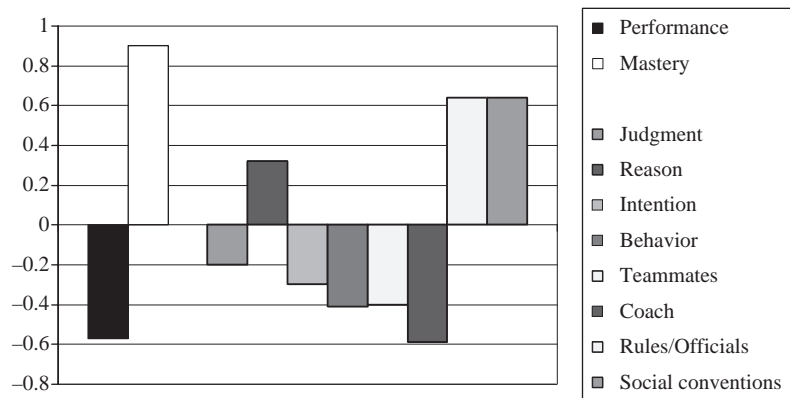

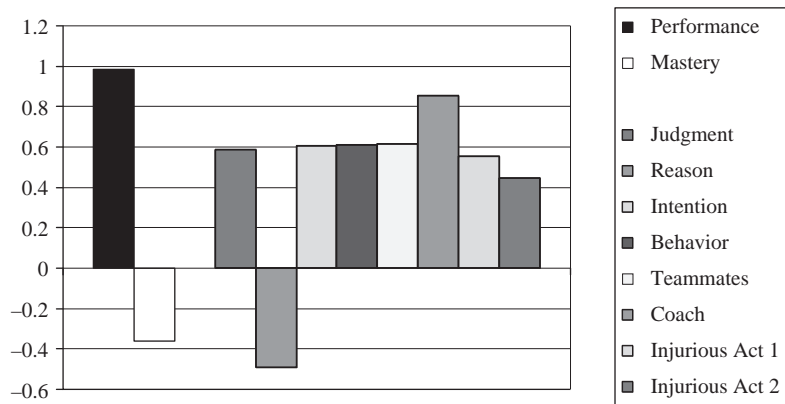*FIGURE 3*   Boys: Low performance, high mastery.

FIGURE 4   Girls: High performance, low mastery.

Gender differences have been found in the motivational climate literature consistently in that young male players react more strongly to performance climate criteria and are more prone to cheating compared with female players: Again, boys cheat more than do girls. However, the important point here is that regardless of gender, a strong coach-created performance motivational climate is related to low moral functioning, inappropriate behavior is judged to be appropriate, and players have lower moral reasoning, have a higher intention to cheat, and actually do cheat more. In addition, both male and female players identified their team moral atmosphere to be more supportive of cheating. These data show clearly that when players are ego or task involved, they think differently, have different ideas about cheating, and report different episodes of intention to cheat and actual cheating. Ego-involved players have more tolerance toward and support cheating to win, even to the extent of supporting injurious acts.

The research findings are quite consistent. When the coach emphasizes criteria of success and failure showing that he or she values winning above all else, athletes indicate that they tolerate cheating more and actually do cheat more than athletes who perceive their coach as emphasizing more mastery criteria of success and failure. This is true for elite athletes, at the professional and Olympic levels, as well as for child and youth sport participants. Clearly, the coach matters, but the major point here is that the coach does not have to endorse cheating per se; rather, it is simply that the coach who emphasizes winning above all else creates a motivational climate that the athletes interpret as being more accepting of cheating behavior.

7. WHAT CAN WE DO ABOUT CHEATING?

Some people deliberately cheat. There are enough cases in sport, in corporate business, and in everyday life demonstrating that some individuals and organizations simply cheat to gain an unfair advantage. There is little that can be done about these systematic cheats except to try to catch and punish them to the full extent possible. Whether that is through drug testing in sport, financial oversight in corporate business, and so on, regulations and oversight are needed to cull out the cheats. But cheating can be much more subtle and unintentional in many cases. It is this latter form of cheating that is addressed in these final remarks.

Competitive sport often places people in conflicting situations, where winning is emphasized and where fair play and justice are sometimes deemphasized. It would be wrong to attribute this to the competitive nature of sport. The research reported in this article suggests that the culprit is not competition in and of itself. Rather, the culprit may be the difference in salience of ego- and task-involving cues in the environment, whether through one's own disposition or through the cues emphasized by one's coach (or parent). It is these perceived criteria of success and failure that induce differential concern for moral action and cheating. A focus on winning may reinforce prejudice (i.e., "us against them") and lead to a more depersonalized view of opponents that, in turn, makes cheating more possible.

How can the development of good sportspersonship and sound moral reasoning be maximized, and how can cheating be prevented? Clearly, one way in which to help is to teach the coach to reinforce the importance of

task-involving achievement criteria in the competitive environment. This is important at all levels, especially in light of the degree of cheating going on in professional and Olympic sports. But this is more important for children's and youth sports. This does not mean that one cannot strive to win; rather, it means that we should apply different criteria to the meaning of winning and losing. By giving feedback to players on effort, hard work, and self-referenced increments in competence, more respect for rules, conventions, officials, opponents, and appropriate moral reasoning and action could be cultivated. One way in which to say this is that there are no problem athletes, just problem coaches (and parents). Coaches must be aware of the criteria of success and failure that they are emphasizing. If task-involving criteria are emphasized, cheating can be reduced and moral reasoning and action can be enhanced.

When focusing on children in the competitive sport experience, it is not just the coach who needs to be aware of the criteria being emphasized; parents are the most important source of information for the children. Indeed, it is often the parents who are most to blame in children's sports. However, coaching programs that address both parents and coaches now exist in most countries. The evidence is accumulating that strategies and instructional practices should be developed to facilitate the coach (and parents) in creating a task-involving coaching environment that is concerned with mastery criteria of success and failure. If there is concern about optimizing the motivation, psychosocial development, and moral functioning and action of children, the impact that teaching and coaching style have on these outcomes cannot be ignored.

However, coaches must realize that simply trying to enhance mastery criteria is often not enough. The ego-involving criteria are emphasized as well. Normative evaluation should not be used for athletes. Comparisons should be made with children's own past performance rather than with the performance of their peers. By giving feedback based on their own past performance and trying to give feedback to athletes privately rather than in an announcement to the whole team, coaches can do their best to maintain the motivation of the athletes and remove an important determinant of cheating in sport.

Will cheating ever be eliminated from sport? That is unlikely in elite sport when the rewards are so high. But in children's sport, attempting to reduce cheating is critical. Playing fair and recognizing justice for all should be a natural part of the sport experience. If the sport experience is the context in which children learn to relate to their peers, not only in sport but also for life, the criteria of success and failure inherent in the sport experience should be a critical concern.

See Also the Following Articles

Fair Treatment and Discrimination in Sport ■ Goal Setting and Achievement Motivation in Sport ■ Intrinsic and Extrinsic Motivation in Sport

Further Reading

Dunn, J. G. H., & Dunn, J. C. (1999). Goal orientations, perception of aggression, and sportsmanship in elite male youth ice hockey players. *Journal of Sport and Exercise Psychology*, *13*, 183–200.

Kohn, A. (1986). *No contest: The case against competition*. Boston: Houghton Mifflin.

Nicholls, J. G. (1989). *The competitive ethos and democratic education*. Cambridge, MA: Harvard University Press.

Ommundsen, Y., Roberts, G. C., Lemyre, P. N., & Treasure, D. (2003). Perceived motivational climate in male youth soccer: Relations to social–moral functioning, sportspersonship, and team norm perceptions. *Psychology of Sport and Exercise*, *4*, 397–414.

Rest, J. R. (1984). The major components of morality. In W. Kurtines, & J. Gewirtz (Eds.), *Morality, moral behaviour, and moral functioning* (pp. 356–429). New York: John Wiley.

Roberts, G. C. (2001). Understanding the dynamics of motivation in physical activity: The influence of achievement goals on motivational processes. In G. C. Roberts (Ed.), *Advances in motivation in sport and exercise* (pp. 1–50). Champaign, IL: Human Kinetics.

Shields, D. L., & Bredemeier, B. J. L. (1995). *Character development and physical activity*. Champaign, IL: Human Kinetics.

Vallerand, R., & Losier, G. F. (1994). Self-determined motivation and sportsmanship orientations: An assessment of their temporal relationship. *Journal of Sport and Exercise Psychology*, *16*, 229–245.

Child Custody

David A. Martindale
Morristown, New Jersey, USA

Jonathan W. Gould
Charlotte, North Carolina, USA

GLOSSARY

best interests of the child doctrine The philosophy, now incorporated into the law in all 50 states, that in making decisions concerning custodial placement and parenting plans, the interests of the children should supersede those of the parents and/or others involved in the dispute.

evaluation of comparative custodial suitability A forensic psychological evaluation the primary focus of which is the relative strengths and deficiencies in the ability of parents to parent the specific children whose custodial placement is at issue.

forensic role Within the context of custody evaluations, a perspective and actions that are consistent with one's obligation to function as an impartial examiner and to gather information that bears upon specific issues before the court.

pendente lite A reference to a matter that one or both parties assert must be addressed by the court while the primary issues in litigation are still pending.

therapeutic role Within the context of custody evaluations, a perspective and actions that are consistent with one's obligation to assist those with whom one is professionally interacting.

This article begins with a historical overview of the manner in which child placement disputes have been adjudicated and a discussion of the best interests of the child standard. The outline of the factors that psychologists endeavor to assess includes a discussion of the methodological problems inherent in this form of assessment. The significant differences between therapeutic work with families and forensic assessment of families are explained. Because performing custody evaluations requires that psychologists work in a collaborative manner with attorneys and adapt to procedures with which mental health professionals are often unfamiliar, the adjustments that must be made are discussed. The article concludes by providing some perspectives on experts offered by two judges.

1. PSYCHOLOGY'S CONTRIBUTION TO UNDERSTANDING BEST INTERESTS

The historical context of custody evaluations begins with Roman law and the presumption of paternal preference. Children were viewed as the property of their fathers. A father had absolute power over his children. He was allowed to place his children into slave labor or sell them for profit. A child's mother had no legal rights. English common law also provided for absolute paternal power. Children were viewed as the property of the father and he had sole discretion over where the children lived, both during the

marriage and after a divorce. Mothers had restricted access to their children after divorce.

Though the reasoning varied, custodial placement decisions based primarily on generally accepted sex roles were the norm until 1970. In 1813, in deciding on the custodial placement of two children, ages 6 and 10, the Supreme Court of Pennsylvania declared: "It appears to us, that considering their tender age, they stand in need of that kind of assistance which can be afforded by none so well as a mother" [*Commonwealth v. Addicks and Lee*, 5 Binney's Rep. 520 (1813), at 521]. At approximately the same time, in Great Britain, what came to be referred to as the tender years doctrine dictated that children younger than age 7 required the care of their mothers. From age 7 on, the custodial responsibility for children would revert back to the father. As the work of fathers began to take them outside the home and as mothers became the primary caretakers, a preference for placement of children with their mothers developed. By the 1920s, a presumption in favor of the mother was clear in most American courts.

Comfort with the notion of a maternal preference was buttressed by Freud's then-popular theory that, among other things, stressed the unique role of mothers in the lives of children. Theories of bonding contributed further to the view that children needed to be with their mothers. Current research has demonstrated that infants form meaningful attachments to both parents by the middle of their first year; however, only recently has such research raised questions concerning the wisdom of a maternal preference.

In the 1960s, several social changes stimulated a reexamination of the maternal preference. These factors included an increase in the number of fathers wishing to maintain an active postdivorce role in the lives of their children, the entry into the workforce of large numbers of mothers, and an increased sensitivity to the discriminatory effects of placement decisions predicated on sex.

In 1970, the U.S. Congress passed the Uniform Marriage and Divorce Act, providing for a best interests of the child standard. The tide had turned from a focus on paternal or maternal preference or rights to a focus on what is best for the child. Simultaneously, psychology offered a revolutionary idea: The concept of the psychological parent was developed. Once society had decided that children, being the most innocent victims of divorce, should be protected as effectively as possible from its emotional consequences, the view emerged that mental health professionals could assist triers of fact in determining what parenting plan was most likely to best serve the interests of the children.

The best interests standard is not without its critics. It has been criticized as poorly defined, as too ambiguous, and as providing fodder for disputes and litigation. Mnookin has argued that the broad judicial discretion provided by the best interests standard leaves judges free to impose their personal biases and beliefs on their judicial determinations. The Supreme Court, in deciding *Troxel v. Granville* (a grandparents' visitation case), made note of the fact that the Washington Superior Court trial judge (whose decision was overturned) had deemed it appropriate to "look back at some personal experiences" [*Troxel v. Granville*, 530 U.S. 57 (2000), at 61].

The challenge is to develop a legal standard that is sensitive both to the social values of American culture and to the needs of the legal system. A legal standard addressing the best interests of the child should "offer effective and useful guidelines, so that similar cases are decided similarly without extralegal factors significantly affecting final dispositions" (Krauss & Sales, 2000, p. 845). A legally effective child custody standard would balance the needs of the family, the needs of society, the evolving social science research, the evolving case law precedents, and the changes in each family configuration while minimizing state intrusions into the functioning of the family.

In many states the criteria to be utilized in evaluating comparative custodial suitability are defined by statute. In states where criteria are not so defined, experts conducting evaluations of comparative custodial fitness bear an obligation to articulate the criteria that they will utilize in performing their assessments.

2. PSYCHOLOGICAL ASSESSMENT: WHAT CAN WE MEASURE?

In 1971, the Supreme Court handed down a decision in *Griggs et al. v. Duke Power Co.* (401 U.S. 424), a case that, on its face, had nothing whatsoever to do with children and custodial placement decisions. The Griggs case involved assessment in the selection, placement, and promotion of personnel. The court ruled that testing procedures must be demonstrably reasonable measures of (or predictors of) job performance. In the years since 1971, forensic psychologists involved in custody work have come to view the Griggs decision as an exhortation, urging us to focus our attention and our assessment efforts on functional abilities that bear directly on the issue before the court. Within the context of custody

work this suggests that we must endeavor to ascertain what attributes, behaviors, attitudes, and skills are reliably associated with effective parenting and confine ourselves to assessing those characteristics as they relate to the children who are the focus of the dispute.

Our problem is that the only assessment instruments with established reliability and validity provide data pertaining only to one or two of the criteria of potential interest to us. An instrument such as the MMPI-2, for example, may provide useful information concerning a test taker's general mental/emotional health, but that leaves much to be assessed about the individual's parenting. LaFortune and Carpenter (1998, p. 222) list the seven most frequently used assessment instruments that "focus on parenting skill and the parent–child relationship" and that are "touted by their authors as helpful in clinical determinations of parental fitness." LaFortune and Carpenter declare: "[T]he validity of these measures is unestablished at best and seriously flawed at worst" (p. 222). Although the hope is expressed that improvements in these instruments may make them useful in the future, the authors conclude that their use at present "cannot be recommended" (p. 222).

Although frequency of use by surveyed mental health professionals is not an acceptable measure of either the reliability or the validity of an assessment instrument, a disturbingly high proportion of custody evaluators use instruments that are psychometrically deficient but that are popular, largely because of successful marketing. Additionally, some clinicians have brought to forensic work psychodiagnostic assessment instruments generally considered useful in health care settings but ill suited to the evidentiary demands of forensic work. In a therapy setting, a psychodiagnostic assessment marks the beginning of an ongoing relationship in the course of which there will be opportunities for subsequent reassessment. In a forensic setting, the report in which one's assessment is described marks the end of a relationship. No opportunities to reassess are provided. Because of this critical difference between clinical assessment and forensic assessment, it cannot be presumed that instruments popular among clinicians are suitable in forensic work.

Consider the following, from the Standards for Educational and Psychological Testing (American Education Research Association *et al.*, 1999): Tests are to be accompanied by documentation that will provide test users with "the information needed to make sound judgments about the nature and quality of the test, the resulting scores, and the interpretations based on the test scores" (p. 67). Elsewhere, the authors of the Standards opine: "The greater the potential impact on test takers, for good or ill, the greater the need to identify and satisfy the relevant standards" (p. 112).

The authors of the Standards have called attention to the fact that "[w]hen directions to examinees, testing conditions, and scoring procedures follow the same detailed procedures, the test is said to be standardized. Without such standardization, the accuracy and comparability of score interpretations would be reduced" (p. 61). Many of those currently performing evaluations of comparative custodial suitability came to this work via clinical psychology, where being helpful is not viewed as an impediment to assessment. Moreland, a former member of the NCS staff and the MMPI-2 restandardization team, observed that "in responding to questions, examiners may inadvertently hint at the nonpathological answer when the test taker is someone whose position they favor (whether it be a custody case, injury case, or whatever)." He added: "[N]either NCS nor the Eyde group has told test givers how to deal with questions from test takers" (e-mail from Kevin Moreland, August 22, 1996).

In the vast majority of custody disputes, it is presumed by all involved that both parents are suitable. The court's task, with advisory input from the mental health professional, is to decide whether, in addition to being capable, the two parents can coparent effectively. If they can, a joint custodial arrangement may be appropriate; if they cannot, one of these two suitable parents must be designated as the primary custodian. If a primary custodian must be designated, then the court must decide which of two presumably suitable parents is more suitable. Psychologists, and those who seek our advisory input in custodial placement disputes, must recognize that we have no way of evaluating the validity of our recommendations because we have no way of knowing what would have occurred if the nonfavored parent had, instead, been selected as the primary custodian. Otto *et al.* (2003) observed that "research in this area will always be constrained by the inability to use true experimental designs to address the most prominent questions related to custody" (p. 203).

The foregoing paragraphs have an unmistakably pessimistic tone. Some might argue that we do not do custody evaluations well and that if we cannot do them well, we should not do them at all. It is reasonable to presume that as increasingly more psychologists become involved in assessing comparative custodial fitness and share with their colleagues their ideas for improving our procedures, we will become increasingly

more proficient. Deciding not to do custody evaluations is not the answer.

3. THE MARRIAGE OF CLINICAL SKILLS AND FORENSIC PRECISION

Most psychologists currently practicing or preparing to practice forensic psychology obtained their original training in clinical psychology. Forensic work is built on solid clinical training and experience; however, to function effectively as a forensic psychologist, the therapeutic mind-set must be altered significantly. It is also important to recognize that there is, in the words of Greenberg and Shuman (1997), an "irreconcilable conflict between therapeutic and forensic roles" (p. 50). A psychologist who is serving or who has served in a therapeutic capacity with any of those involved in a custody dispute should not function as the evaluator. Similarly, an evaluator should not offer therapy or counseling services to those whom he or she has evaluated.

In a forensic setting we seek answers to questions that have been posed (either directly or implicitly) by the court. We are not attempting to determine the etiology of a litigant's problems, nor are we attempting to set therapeutic goals and formulate treatment plans. Instead of communicating warmth, we communicate skepticism. Although that may sound harsh, it is important that litigants be reminded that information offered by them will not simply be written down in our notes and accepted at face value. Reasonable attempts will be made by us to verify information given to us. Section VI.F of the *Specialty Guidelines for Forensic Psychologists* (Committee on Ethical Guidelines for Forensic Psychologists, 1991) reminds forensic psychologists that "[w]here circumstances reasonably permit, forensic psychologists seek to obtain independent and personal verification of data relied upon." Ceci and Hembrooke (1998, p. 4) have expressed the view that "the uniquely human qualities that compel us to want to help are the very qualities that can make us poor experts."

Some of the most serious problems that forensic psychologists create for themselves stem from an inability or unwillingness to control the impulse to think like and act like "helpers" when they are obligated, by virtue of their forensic task, to function as examiners. Shuman (1993, p. 298) has opined that the techniques customarily employed by mental health professionals to establish a therapeutic alliance "are entirely inappropriate in

a court-ordered examination." Pruett and Solnit (1998, p. 126) express the view that "empathic skills must be constrained in a manner unusual for clinical practice but essential for fair, ethical, evaluative interaction."

The prevailing presumption in clinical work is that those who appear before us honestly share with us their perceptions of themselves and those with whom they interact. There is a tendency to presume, further, that discrepancies between a client's descriptions and objective reality are attributable, primarily, to the operation of various perception- and memory-distorting phenomena of which the client is not consciously aware. Only limited thought is given to the possibility of deliberate, calculated lying.

Individuals being examined within the context of custody disputes are strongly motivated to present themselves as being better adjusted than they know themselves to be. An examination that might be deemed adequate in a clinical context might be viewed as inadequate in a forensic context. Regardless of the assessment instruments utilized, a forensic examiner must operate more like an investigator than a sympathetic and supportive listener. In particular, custody evaluators cannot be passive recipients of offered information. We must use clinical skills to elicit information and actively seek, both from the litigants and from appropriate collaterals, any information that might reasonably be viewed as pertinent. Data collected by Ackerman and Ackerman suggest that approximately 28% of the evaluative time expended by those responding to a survey was devoted to information verification (obtaining information from documents, disinterested collateral sources, and other nonparties). Bow and Quinnell found that 98% of evaluators review documents relevant to the case, 86% interview therapists, 78% interview teachers, and 52% interview physicians and/or pediatricians.

4. DANCING WITH ATTORNEYS WITHOUT TRIPPING

In the field of custody work, most of the missteps that occur in interactions between forensic psychologists and attorneys relate to conflicting obligations. Such conflicts are easy to understand when psychologists are endeavoring to function as impartial examiners. Clashes are often unanticipated when psychologists have been retained by attorneys to function as their consultants or to offer expert testimony. Even when functioning in the consultant role, however, psychologists' ethical

obligations differ from those of the attorneys who have retained them and choreography can still be a challenge.

It is not uncommon for judges and attorneys to ask that a psychologist functioning as impartial evaluator alter his or her customary data collection sequence in the name of expediency or offer interim recommendations to assist the court in resolving a *pendente lite* issue. With regard to data collection, there is a logical progression to the information-gathering aspect of a custodial fitness evaluation, and although some modifications are acceptable, others are not. Data obtained through the administration to the parties of psychological tests and information secured by means of various questionnaires inevitably suggest areas of inquiry. It is inadvisable to conclude one's evaluative sessions with the litigants without having in hand test data and the questionnaires.

Standard 9.01 (a) of the psychologists' ethics code (American Psychological Association, 2002) declares: "Psychologists base the opinions contained in their recommendations, reports, and diagnostic or evaluative statements, including forensic testimony, on information and techniques sufficient to substantiate their findings." In the absence of sufficient information, recommendations should not be offered. All involved should also be mindful of the fact that parenting plans set in place in response to specific and immediate concerns alter family dynamics in ways that are not readily predictable and that may not be in the best interests of the children.

Court appointment does not guarantee evaluator competence. Neither does competence assure findings and recommendations with which all involved will be pleased. Where an expert's opinions have not been formulated through the utilization of appropriate procedures and are not supported by reliable data, exposing these deficiencies is a critically important function of an opposing expert.

The *Specialty Guidelines for Forensic Psychologists* (Committee on Ethical Guidelines for Forensic Psychologists, 1991) closes with a reminder (contained in section VII.F) that "[f]orensic psychologists are aware that their essential role as expert to the court is to assist the trier of fact to understand the evidence or to determine a fact in issue." Psychologists retained as consultants to attorneys or retained to offer expert testimony are, despite their collaborative relationships with those who have employed them, obligated to promote understanding and to decline participation in partisan attempts to distort or misrepresent evidence. Just as those whom we evaluate should be provided with information concerning our policies and procedures, so, too,

should attorneys who retain us be provided with such information. Information provided in advance prevents problems.

5. FOLLOWING THE RAINBOW

Standard 2.01 (c) of the psychologists' ethics code states, in part: "Psychologists planning to provide services ... new to them undertake relevant education, training, supervised experience, consultation, or study." Section 2.01 (e) states, in part: "In those emerging areas in which generally recognized standards for preparatory training do not yet exist, psychologists nevertheless take reasonable steps to ensure the competence of their work" (p. 1064). In 2002, the American Psychological Association formally designated forensic psychology as a specialty. The obligation of psychologists trained in clinical psychology to professionally prepare themselves for forensic work if they plan to offer forensic psychological services is clear yet occasionally ignored.

The performance of an evaluation of comparative custodial suitability is, by definition, a forensic psychological endeavor. In our view, such evaluations should be guided by the forensic model. Earlier, we alluded to the Supreme Court's expressed disapproval of a trial court judge's decision to "look back at some personal experiences" [*Troxel v. Granville*, 530 U.S. 57 (2000), at 61] in formulating his opinion in a visitation dispute. The opinion formulation process to which the judge made direct reference is, more often, unarticulated. There is reason for concern that psychologists, whose services have been sought because of their presumed expertise, occasionally reflect on personal experiences, values, and beliefs in formulating their opinions. What makes an opinion an expert opinion is not the professional credentials of the individual whose lips form the words or whose keyboard is used to type them. What makes an opinion an expert opinion is the manner in which the opinion was formulated. Professional credentials are necessary but not sufficient.

An expert opinion is formulated utilizing the accumulated knowledge and the recognized procedures of the expert's field. If the accumulated knowledge of the expert's field was not utilized, the opinion expressed is not an expert opinion. It is a personal opinion, albeit one being expressed by an expert. Judge Alex Kozinski has astutely observed that the task facing triers of fact, when expert testimony has been offered, "is to analyze not what the experts say, but what basis they have for saying

it" [*Daubert v. Merrell Dow Pharmaceuticals, Inc.* (on remand), 43 F.3d. 1311 (9th Cir. 1995), at 1316]. The Honorable Stephen Hjelt, the presiding Administrative Law Judge for the California Office of Administrative Hearings in San Diego, in what might best be described as an open letter to psychologists, opined that our "profession has strong roots as a discipline that has a foundation in the scientific method. However, some of you simply stopped using it" (Hjelt, 2000, p. 12). We close by urging psychologists, attorneys, and judges to consider the words of judges Kozinski and Hjelt. Psychologists should use their training in methodology, cross-examining attorneys should demand that testifying experts articulate the bases for their opinions, and judges should consider the data that form the basis for expert opinions and not be gullible consumers of recommendations offered by experts.

See Also the Following Articles

Diverse Cultures, Dealing with Children and Families from ■ Eyewitness Identification ■ Forensic Mental Health Assessment

Further Reading

Ackerman, M. J., & Ackerman, M. C. (1997). Custody evaluation practices: A survey of experienced professionals (revisited). *Professional Psychology: Research and Practice*, 28, 137–145.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060–1073.

Bow, J. N., & Quinnell, F. A. (2001). Psychologists' current practices and procedures in child custody evaluations: Five years after American Psychological Association Guidelines. *Professional Psychology: Research and Practice*, 32, 261–268.

Ceci, S. J., & Hembrooke, H. (1998). *Expert witnesses in child abuse cases: What can and should be said in court.* Washington, DC: American Psychological Association.

Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, 15(6), 655–665.

Coons, J., Mnookin, R., & Sugarman, S. (1993). Deciding what's best for children. *Notre Dame Journal of Law and Public Policy*, 7, 465–490.

Eyde, L. D., Robertson, G. J., Krug, S. E., Moreland, K. L., Robertston, A. G., Shewan, C. M., Harrison, P. L., Porch, B. E.,

Hammer, A. L., & Primoff, E. S. (1993). *Responsible test use: Case studies for assessing human behavior.* Washington, DC: American Psychological Association. [The authors are also refered to, collectively, as the Test user Training Work Group of the Joint Committee on Testing Practices.]

Greenberg, S. A., & Shuman, D. W. (1997). Irreconcilable conflict between therapeutic and forensic roles. *Professional Psychology: Research and Practice*, 28(1), 50–57.

Hjelt, S. (2000). Professional psychology: A view from the bench. *Register Report*, 26(1), 8–13.

Kelly, J. B., & Lamb, M. E. (2000). Using child development research to make appropriate custody and access decisions for young children. *Family & Conciliation Courts Review*, 38(3), 297–311.

Krauss, D. A., & Sales, B. D. (2000). Legal standards, expertise, and experts in the resolution of contested child custody cases. *Psychology, Public Policy, & Law*, 6(4), 843–879.

LaFortune, K. A., & Carpenter, B. N. (1998). Custody evaluations: A survey of mental health professionals. *Behavioral Sciences and the Law*, 16, 207–224.

Lamb, M. E., & Kelly, J. B. (2001). Using empirical literature to guide the development of parenting plans for young children: A rejoinder to Solomon and Biringen. *Family Court Review*, 39, 365–371.

Martindale, D. A., & Gould, J. W. (2004). The forensic model: Ethics and scientific methodology applied to custody evaluations. *Journal of Child Custody: Research, Issues, and Practices*, 1, 1–22.

Mnookin, R. (1975). Child custody adjudication and judicial function in the face of indeterminacy. *Law and Contemporary Problems*, 39, 226–293.

Otto, R. K., Buffington-Vollum, J. K., and Edens, J. F. (2003). Child custody evaluation. In A. M. Goldstein (Ed.), *Handbook of psychology: Vol. 11, Forensic Psychology* (pp. 179–208). New York: Wiley.

Pruett, K. D., & Solnit, A. J. (1998). Psychological and ethical considerations in the preparation of the mental health professional as expert witness. In S. J. Ceci, & H. Hembroke (Eds.), *Expert witnesses in child abuse cases: What can and should be said in court* (pp. 123–136). Washington, DC: American Psychological Association.

Shuman, D. W. (1993). The use of empathy in forensic examinations. *Ethics and Behavior*, 3, 223–302.

Warshak, R. A. (2000). Blanket restrictions: Overnight contact between parents and young children. *Family and Conciliation Courts Review*, 38(4), 422–445.

Whiteside, M. F. (1998). Custody for children age 5 and younger. *Family & Conciliation Courts Review*, 36(4), 479–502.

Child Development and Culture

Cigdem Kagitcibasi

Koc University, Istanbul, Turkey

GLOSSARY

cultures of relatedness Collectivistic societies where interpersonal distances are small between connected selves.

heritability The portion of variability in a trait that is passed to offspring via genes; high heritability is often understood to mean high genetic causality.

mechanistic model A behaviorist approach that interprets "environment" in a very limited sense as proximal stimuli.

organismic model An approach that emphasizes biologically based maturation, ignoring the context of development.

scaffolding Temporary support and guidance in problem solving.

Human development occurs in a cultural context. Obvious as this statement may be, its realization in developmental science is far from adequate. Most mainstream research and theory that derives from the Western (particularly American) academic centers and informs the field tends to approach human development in and of its own, without much attention to its cultural aspects. This state of affairs has implications for how child development is construed, how it is assessed, and what is done (if anything) to enhance it. This article presents the current knowledge regarding the interface of culture and child development, with the main processes and dynamics involved, and also provides a brief overview of the scholarship on the subject. The main theoretical conceptualizations regarding child development, as well as their implications in significant spheres of applied fields, are reviewed.

1. STUDY OF CHILD DEVELOPMENT IN PSYCHOLOGY

A main issue in the psychological study of child development is the scarcity of a contextual approach. The prominence of the mechanistic and organismic models in developmental psychology and the stress on the individual as the unit of analysis have worked as deterrents to contextual conceptualizations of child development. The mechanistic model, used by behaviorism, has construed "environment" in a very limited sense as proximal stimuli. The organismic model has stressed biologically based maturation, ignoring the context of development. These perspectives are of both historical and contemporary influence. Recent advances in cognitive neuroscience, genetic research, and the like have further strengthened the "biological shift." The focus has been on the individual organism as a product of these two major influences—more recently, as a combination of the two. Thus, a vast

majority of current psychological research on child development is conducted in the laboratory and does not take the larger cultural context into account.

Notwithstanding this state of affairs, there is also a rich tradition of research and thinking in child development, particularly in anthropology, that takes culture seriously. Introducing culture into child development has been done, starting mainly during the second half of the 20th century, first in anthropology and then in ecological, cross-cultural, and cultural psychology. Although this work remained rather marginal for some decades, it is beginning to be better recognized, to increase in volume, and to leave its mark on mainstream research and conceptualizations in developmental psychology. One reason for this changing orientation is the fact that research informed by cultural context is now being done in modern nations, including those in the Western world, rather than in isolated small human groups, as was the case with classic anthropological work. So, it is seen as more relevant by psychologists. A second reason is the fact that Western societies, especially the United States where there is the greatest amount of research production, have become multicultural societies. The growing cultural diversity has again made culture more relevant in the psychological study of development.

1.1. Study of Child Development Within Culture

Research on child development that takes culture seriously is being done within two different paradigms. One of these, known as cross-cultural psychology, uses a comparative approach in which theory testing is of high priority. Cross-cultural comparative methodology works as a corrective to the rather ethnocentric assumption of the pan-human generality of findings obtained in unicultural studies. The view here is that the findings may in fact have universal generality, but this has to be demonstrated rather than assumed. Indeed, the theoretical basis for a comparative methodology is the "universalism" of psychological processes, although their manifestations in behavior differ due to cultural factors. The methods/instruments used are similar to those used in unicultural psychological studies of development; however, the issue of equivalence in measurement emerges as much more significant in cross-cultural methodology.

The other paradigm is cultural psychology, which claims that cultural context gives meaning; thus, human behavior is context specific and should be studied within culture, not comparatively. For example, topics such as the cultural organization of human development within family and society and the cultural construction of intelligence are studied. This is akin to an anthropological perspective stressing the uniqueness of each cultural context. Accordingly, research tends to be based on observations within context and descriptive qualitative analysis rather than on experimentation and standardized measures with a comparative orientation that are typical in cross-cultural research. For example, observational studies have been conducted extensively within the tradition of "everyday cognition" or "informal learning" by children of culturally relevant skills such as tailoring, woodcarving, weaving, pottery, and other handicrafts as well as street mathematics.

The two paradigms have also been called etic and emic, in analogy with phonetics and phonemics in linguistics. Phonetics studies general aspects of sounds in languages, whereas phonemics studies sounds in a particular language. The different perspectives underlying cultural and cross-cultural psychology and the etic and emic approaches are posited as conflicting; however, they both contribute to our understanding. They both have emerged as correctives to the "culture-blind" orientation of mainstream developmental psychology's organismic and mechanistic paradigms. However, they both entail some risks when taken to an extreme. The comparative cross-cultural approach can claim "false universals," whereas the cultural approach can assume "false uniqueness" (of each cultural context). The challenge is to combine the benefits of each in understanding the unity and diversity within the dynamics of culture and child development.

1.2. Life Stages and Culture

Over the past few decades, the study of human development has moved from a focus on childhood to a focus on the life span. The span of development is seen to be continuous and to cover the whole spectrum, from prenatal period to old age and death. This approach renders culture even more relevant because cultural differences emerge quite significantly during the various phases of the life span. Thus, although there tends to be a great deal of commonality in the meaning attributed to, and the socially described role of the growing human during, infancy and early childhood, cultural variations start and increase progressively during the subsequent stages of middle and later childhood, adolescence, young and later adulthood, and old age. For example, early anthropological work showed that adolescence

and gender roles during adolescence in some preliterate societies are quite different from those in Western technological societies. Even when less diverse socioeconomic cultural contexts are compared, such as more affluent/urban and less affluent/rural ones, the roles of both adolescents and children in family and society are found to differ. In particular, a much greater proportion of children and adolescents work, carrying out household chores or family business, and are economically productive in less affluent/rural and more traditional society. For example, the fact that children in Africa spend less time in "play" and more time in "work," as compared with children in the United States, can be attributed to many factors such as the necessities of less affluent lifestyles but also the cultural meaning of childhood and children's work. Nsamenang claimed that child work in West Africa is a social learning process for adult roles and, therefore, an inherent part of development. From a Western human rights perspective, however, child work can be seen as child "labor"—and, therefore, child abuse—especially if it interferes with formal education. This is a case in point for the various cultural meanings accompanying different lifestyles that can be a source of debate or even conflict.

Thus, cultural meanings associated with different life stages reflect lifestyles, social norms, and social values. Old age is another life stage that is defined and understood variously. Whereas old people in Western technological societies tend to lose social status and experience some isolation from the community of young people, old people in many traditional societies remain very much integrated within family and community and are respected for their life experience, knowledge, and wisdom. They also tend to be more active in important domestic and communal roles involving imparting of traditions to the young, moderating dispute settlements, and the like.

Thus, culturally informed life span perspectives show that age does not have the same meaning across space and time. What is universal is possibly the tendency of all cultures to use age as a basis for defining the individual and assigning him or her with certain roles, responsibilities and privileges. With this in mind, this article focuses on the life stage of childhood in the cultural context.

2. DEVELOPMENT IN CONTEXT

The context of child development includes numerous levels of influences, all interrelated with and embedded within one another. One way of conceptualizing

contextual variables is in terms of their degree of comprehensiveness as encompassing systems. This ecological perspective has been influential in the study of child development in the cultural context. For example, Bronfenbrenner differentiated among four levels, increasing in complexity and comprehensiveness from micro-, to meso-, to exo-, to macrosystems. Determining which of these ecological systems is to figure in the study of child development at any particular time is an empirical issue. An important component of the macrosystem of the ecology of child development is societal values regarding children that are reflected in parental beliefs, values, and child-rearing orientations.

The recognition of the significant implications of the ecological perspective for child-rearing orientations and the resultant child outcomes have led to concepts such as "ethnotheories" or "naive theories" (e.g., LeVine's theory of parental goals). Super and Harkness proposed the concept of the "developmental niche," consisting of the physical and social setting, child care and child rearing, and the psychology of the caretaker, to explain the cultural structuring of child development. Similarly, the "ecocultural niche" (suggested by Weisner), and the "developmental microniche" (suggested by Worthman) have been proposed as organizing concepts to model the relationships of biology, behavior, and culture in shaping human development. Other theoretical models have also entailed an ecological/cultural contextual framework, such as Berry's "ecocultural framework," further combined with ecological systems theory by Georgas and with the developmental niche by Dasen.

An ecological perspective also calls for situating parental beliefs and child-rearing orientations within the socioeconomic and social structural context. This is important to understand why certain belief systems show systematic variation across different parental populations. The answer often lies in the underlying functional relations. Ecological perspectives have also figured in other analyses. For example, bringing the temporal dimension into an ecological perspective in forming a sociological life events approach, Elder studied intergenerational relations and adolescent adjustment and development at various historical periods. Similarly, concepts of developmental pathways and "canalization" over an epigenetic landscape were used by Waddington to analyze variations and self-stabilizing constraints that restrain the variations in the organism–environment system as it changes and evolves over time.

These conceptualizations aim to deal systematically with the myriad spatial and temporal environmental influences that provide developmental pathways for

the individual. Nevertheless, the individual not only is acted on but also acts on the environment so as to create its own developmental pathway that is unique yet also has common characteristics with others. This view is stressed particularly by European-based theories of individualization.

2.1. The Value of Children and Family as Developmental Niche

Study of parental values and child-rearing orientations can serve as an example of an ecological approach to development in the cultural context. Values attributed to children and how childhood is conceptualized provide important clues as to the place of children in the family and society. A cross-cultural study on the Value of Children (VOC), conducted during the 1970s and currently being replicated in several countries, points in particular to the economic/utilitarian and psychological values attributed to children. The economic/utilitarian value is stressed in sociocultural contexts where children's material contributions to the family are significant. This often takes the form of contributing to family income while young and providing elderly parents with old-age security during later years. Psychological VOC, on the other hand, is stressed in contexts where children entail more economic costs than economic assets. These dynamic relations among ecological factors, family/parental values, and child outcomes are dealt with by another ecological theory, namely the "family change model" (proposed by Kagitcibasi).

In sociocultural contexts of low affluence, rural/agrarian/low-socioeconomic status standing where children in fact contribute to the family's material well-being, the economic/utilitarian VOC is salient, child work is common both in and out of the house, and the family is characterized by intergenerational material interdependencies. This is the family model of "interdependence." In contrast, in more affluent contexts where children are in school and are costly to raise, the psychological VOC is salient, child work is negligible, and intergenerational material interdependencies are weak. This is the family model of "independence." The interdependent pattern is more prevalent in collectivistic cultures or "cultures of relatedness" with closely knit human relations. The independent pattern is more typical of the individualistic societies or "cultures of separateness." Thus, there is a general correspondence among societal conceptions of childhood, parental values, family patterns, and children's actual lifestyles.

In the family model of interdependence, there is obedience-oriented child rearing that does not promote the development of autonomy. This is because the independence of children is not valued and might even be seen as a threat to the family livelihood because independent offspring may look after their own self-interests rather than those of the family. Intergenerational interdependence is manifested through the family life cycle first as the dependence of the children on their parents, with this dependence being reversed later on as the dependence of the elderly parents on their adult offspring for their livelihood.

A contrasting pattern is seen in the family model of independence, characteristic of the Western (mostly American) middle-class nuclear family, at least in professed ideals. Here, child rearing is oriented toward engendering self-reliance and autonomy, and independence is also valued between generations, where objective social welfare systems and affluence render family interdependence unnecessary or even dysfunctional. Individuation–separation is considered a requisite for healthy human development and is also reinforced by (popular) psychological teaching.

It is commonly assumed (e.g., by modernization theory) that there is a shift from the model of interdependence to the model of independence with urbanization and economic development. However, recent research shows that the changes in family interaction patterns often do not follow such a linear route. In collectivistic cultures of relatedness where interpersonal distances are small between connected selves, individuals and families do not shift toward individualistic separateness with urbanization and economic development. Connectedness continues, although mainly in the realm of psychological interdependencies, while material interdependencies weaken with changing lifestyles. This leads to a third pattern proposed by Kagitcibasi, the "family model of psychological interdependence." In some ways, it is different from both the traditional (rural) family characterized by total interdependence and the individualistic (urban) family characterized by independence.

Child rearing involves control rather than permissiveness because the goal is not individualistic separation. However, autonomy also emerges in child rearing, mainly for two reasons. First, autonomy of the child is no longer seen as a danger because elderly parents are no longer dependent on their grown-up offspring for their material livelihood. Second, autonomy of the growing person is functional in urban lifestyles where, for example, decision making (rather than obedience) is required in school and in more specialized jobs.

2.2. Development of the Self

Thus, the family is the developmental niche for the self. Studying the family provides insights into understanding the self, in particular how and why the different types of self develop. For example, a distinction is made between the independent self and the interdependent self, and the cognitive and behavioral concomitants or consequences of these two types of self are examined. Thus, current work on culture and self focuses on cross-cultural variability, particularly along the dimension of independence–interdependence. However, why and how these different types of self develop is not well understood except in reference to macrosocietal characteristics such as individualism and collectivism.

Family is the significant mediating variable here that sheds light on the functional underpinnings of self development. In the family model of interdependence the "related self" develops and is interdependent with others. In the family model of independence, the "separate self" develops and is independent. The third family model of psychological interdependence is more complex because it entails both autonomy and connectedness in child rearing, with the resultant "autonomous-related self" (proposed by Kagitcibasi). The coexistence of autonomy with connectedness is important because the two are often seen to be antithetical in psychology. Under the influence of psychoanalytic thinking, particularly object relations theory and the separation–individuation hypothesis, separation is considered to be a requisite of autonomy in human development. Therefore, the implication is that connected selves cannot be adequately autonomous. This is a central debate, particularly in understanding adolescence. A psychoanalytically oriented individualistic perspective, such as that proposed by Steinberg and Silverberg, considers adolescence to be a "second separation–individuation process" where detachment from parents is seen as a requisite for the development of autonomy. Others, such as Ryan and colleagues, suggest that individuation during adolescence is facilitated not by detachment but rather by attachment. Recently, there also has been some recognition of the compatibility of relatedness and autonomy in the adolescent and adult attachment literature. The autonomous-related self challenges the individualistic assumption. It also goes beyond the dualistic conceptualization of the independent self versus the interdependent self that is prevalent in cross-cultural psychology.

Current research provides support for the compatibility of autonomy and relatedness, indeed pointing to their combination as being a psychologically more healthy state that satisfies the two basic human needs for autonomy and connection (merging). The family model of psychological interdependence also finds research support in both the Western and non-Western contexts. For example, a more positive relationship is found between autonomy and relatedness than between autonomy and separateness in both Korean and American samples; positive links, rather than negative links, are found between relatedness to parents and autonomy in adolescents in the United States; combined autonomy and control orientation is noted among Chinese and Korean parents; parental autonomy goals do not imply separateness, and achievement values are associated with parental collectivism (rather than individualism), among Turkish parents in Germany; family interdependencies coexist with some individualistic values in Hong Kong; and Chinese and Chinese American parents are found to endorse both relatedness and autonomy, together with high control of and closeness with their children.

Parental orientations are of crucial importance in leading to diverse developmental outcomes. Systematic variations are noted even in parents' orientations to infants, providing evidence that the various developmental pathways and their combinations may have their roots all the way back to infancy and are reinforced throughout the life span. Such research, informed by cultural and cross-cultural perspectives, opens up new vistas in the development of the self that can shed light on the interface of culture, parenting, and the individual through time.

3. DEVELOPMENT OF COMPETENCE

Child rearing is goal oriented, although this is mostly not made explicit. Often, the goal is competence in the sense that socialization, by definition, implies becoming a competent member of a social group. Competence in this perspective refers to what is culturally valued, thereby showing variation across cultures. This sphere of child development is the meeting ground for the development of the self and cognitive development.

In more traditional contexts, particularly with closely knit human ties and interdependent family systems, cultural conceptualizations of cognitive competence include a strong social component. Because much research from Africa has pointed to this, it has been called "African social intelligence"; however, it is much more widespread, particularly in traditional

collectivistic cultures of relatedness with closely knit human ties. It reflects cultural valuing of social responsibility and sensitivity to others' needs and expectations. Obedience orientation in child rearing, discussed earlier, is also involved here because it engenders dependent children who grow up to be loyal adult offspring. Independence and separation are not socialization goals.

Research also shows that, particularly in nonindustrialized rural/agrarian contexts of low levels of affluence, children develop practical skills that serve the family and reflect material dependence on children's work. For example, from early on, children do household chores, take care of animals and babies, and even cook food for the family—tasks in which most urban middle-class children would perform poorly. However, the same children often do poorly in simple cognitive tasks with which urban middle-class children have no difficulty. Clearly, children's cognitive competence in culturally valued domains gets promoted, whereas development in other domains lags behind if it is recognized at all. For example, Serpell showed that "folk" conceptions of intelligence in a Zambian village differed significantly from what is measured by intelligence tests, even culturally sensitive ones. Dasen found in Ivory Coast (West Africa) that intelligence involved mainly social and practical skills (e.g., "being good with hands").

3.1. Environmental Change

An issue in the development of competence in cultural context is social change. The channeling of competence toward culturally valued domains is adaptive so long as there is stability in lifestyles and societal demands. A problem emerges, however, in contexts of social change when adaptive mechanisms get challenged by modifications in lifestyles. Such modifications often accompany social structural and economic changes, especially urbanization and migration. For example, rural-to-urban mobility in developing countries is of immense proportions, also feeding into international migration. Although the young population (10–19 years of age) in developing countries was mainly rural in 1990 (approximately 600 million rural vs 300 million urban), the rural and urban populations of youth are expected to become equal before 2015. Thus, this significant global human mobility calls for a better understanding of what is adaptive and what is maladaptive in the cultural construction of children's developmental niches in contexts of social change.

For example, research on ethnic minorities in North America and Europe, mostly migrants from other parts of the world, points to misfits between parental beliefs and values and to new environmental demands emerging from urban living conditions, particularly schools. In the face of social change and mobility, the question of whether there is an optimal fit between (traditional) child-rearing orientations and children's developmental trajectories becomes relevant. Such a question challenges cultural relativism, which assumes that what is culturally valued is by definition the right channel of child socialization. The matter turns into an empirical issue open to investigation by research. Such research calls for the possibility of using common (minimal) standards of child development and competence. It also points to the potential of applied work involving intervention.

3.2. Cognitive Development: Piaget and Vygotsky

A great deal of developmental research in the cultural context has focused on cognitive development. A general conclusion is that basic cognitive processes appear to have cross-cultural similarity; however, their manifestation in behavior shows variability given the diversities in cultural/environmental factors that may serve as constraints or reinforcers. The distinction between capacity and performance comes to the fore, as does the issue of culturally sensitive assessment of cognitive competence.

Most theory in psychology, and also in developmental psychology, claims universality, particularly with the weight of the organismic perspective mentioned previously. This is true for Piagetian theory, the generality of which has been subjected to a great deal of cross-cultural empirical testing with variable results. The most important cross-cultural support for the theory is given to the sequencing of stages of cognitive development. Although the particular ages at which the stages or substages are attained, or whether they are attained, show cross-cultural variability, reversal of stage sequence is never found. Beyond this, the first stage of sensorimotor intelligence is found to be a "strong" universal, shown in all studies with infants. The preoperational period is also considered to be a candidate for a strong universal. Even during these early periods, however, differences emerge, for example, in terms of whether objects or social stimuli are handled or attended to more (apparently due to the different types of stimuli, language use, and maternal responsiveness to which infants and young children are subjected in different cultures).

During the period of concrete operations, cultural differences become more important such that in some cultures, some children fail to engage in concrete

operations even when they reach adolescence. Dasen suggested that the qualitative characteristics of concrete operational development (reasoning types) are universal but that the rate of development in certain domains, particularly quantification and spatial reasoning, is influenced by environmental factors. Finally, individuals in many societies never achieve the formal operational stage of cognitive development. Schooling is the main factor that plays a role in whether formal operations are mastered. Actually, this factor is not confined to the nonindustrialized, non-Western contexts. Not all adults in Western technological society achieve this type of thinking. When confronted with this research evidence, Piaget accepted the role of experience and culture for the development of formal operational thinking.

Vygotsky's sociocultural theory of development takes human development out of the confines of the individual and into the cultural and historical realms. As such, it has been important for contextualist perspectives, particularly those of cultural psychologists. Vygotsky stressed the role played by culture, language, and the child's "zone of proximal development" in cognitive development. This is not a stage theory, but the evolution of speech through "social," "egocentric," and "inner" speech periods is sequential. Research on infant–mother interaction across several cultural groups, and informed by Vygotskian thinking, pointed to cultural differences, particularly in various types of maternal responsiveness and speech. Later, the child is aided by the adult through verbal interaction and instruction to reach beyond the current level of competence. The theory that proposes a continuous process of interaction between persons and sociocultural factors is used as a framework in "everyday learning" or "apprenticeship" situations involving "scaffolding" (i.e., temporary support and guidance in problem solving) by an adult. However, it is also criticized as vague and not conducive to measurement or hypothesis testing. As is also apparent from this glimpse of the two most significant theories of cognitive development, diversity and commonality across cultures challenge developmental theory.

4. NATURE–NURTURE DEBATE AND PLASTICITY

Cross-cultural research can help to provide insights into some of the basic issues in child development. This is mainly because a cross-cultural developmental approach uncovers a greater range of variation than does a single-culture study. With increased coverage of diversity, a wider perspective emerges according to which what is typical and what is atypical might need to be redefined. It also becomes more possible to distinguish between biological and environmental influences. That is, the greater the commonality found across varied cultural contexts, the greater the likelihood of a biological basis for a developmental process, whereas a finding of increased diversity across cultures implies environmental causality.

The perennial nature–nurture debate is a case in point. Although there is much greater sophistication in the treatment of the issue today compared with the earlier, rather simplistic comparisons, the issue stands unresolved. In several spheres, including temperament, personality, intelligence, and gender identity, the nature and malleability of child development can be better understood if cultural context is taken into consideration. "Heritability" refers to the portion of variability in a trait that results from heredity. As such, high heritability is often understood to mean high genetic causality and, thus, a negligible role for the environment. Yet even with extremely high heritability, the environment can still have a powerful influence. For example, a surge in the heights of second-generation Japanese children raised in the United States was reported. The heritability in this group was more than .90, yet the American-reared sons were substantially taller than they would have been if they had been reared in Japan. Similarly, it is reported that the heights of young adult males in Japan increased by approximately 3½ inches since World War II. If a trait as highly heritable as height can show such fluctuations in relatively short time periods and across societies, other traits with lower levels of heritability, such as intelligence and personality, can change as well.

A currently popular view stressing biological determination is Scarr's genotype–phenotype theory, which holds that each individual picks and chooses from a range of environments those that best fit his or her genetic potential/orientations. Thus, the theory claims that genetic endowment creates the environments that individuals choose for themselves. When cultural context is taken seriously, however, the limitations of the theory become obvious. The theory assumes that people have equal opportunity or options from which to choose, yet a great many people in the world are not in a position to choose much at all. Environmental opportunities vary greatly; particularly in deprived environments, the genetic human potentials might not be realized. Thus, heritability for intelligence, for example, would be expected to be lower in deprived conditions than the .50 estimated for the general population

and would be expected to be higher in "good" environments than in "poor" ones because the former provide the resources for the biological potential to be realized.

Indeed a great deal of research points both to lower levels of measured intellectual performance and to other indicators of poor competence such as school failure in socioeconomically deprived environments. Other evidence regarding the powerful influence of the context of development and the plasticity of human development comes from studies of adoption after privation and the effects of schooling.

An early study reported massive intelligence quotient (IQ) gains in children born to low-IQ mothers after they were given up for adoption into environmentally enriched middle-class homes. Although children's IQs still correlated with those of their biological mothers, indicating genetic influence, they were much higher than the mothers' IQs and even higher than those of the general population. This study was criticized for its methodology, but its findings were corroborated by later research. For example, recent work with Romanian adoptees from orphanages after the fall of the Ceaucescu regime points to the crucial importance of both early environmental deprivation and plasticity of human development. It was found that severely malnourished children (below the third percentile in head circumference, body size, and developmental level) recovered to normal development by 4 years of age if they were adopted into English homes before 6 months of age. Those who were adopted later, however, did not fare as well.

There is debate concerning the long-term effects of early deprivation. Some consider the deleterious effects to be critical and long-lasting, even referring to the first 2 years as possibly a "critical period" in human development. Others are less certain due to the limited capacity of infants to process their experiences cognitively and/or due to plasticity, particularly with regard to catching up in physical growth. Nevertheless, there appears to be consensus that whatever the genetic potential, environment, particularly during the early years, has an important impact on human development and that even quite high values of heritability still leave plenty of room for environmental enhancement or reduction of human potential.

5. IMPLICATIONS FOR APPLICATION

Beyond "saving" children from deprivation through adoption, there are other more common implications for application of the significant cultural/environmental impact on human development. Some examples of implications for application, involving advocacy and program- and policy-relevant research, would help to put the issues and their possible solutions into perspective.

5.1. Developmental Norms

Growth norms for infants and children have been established by pediatricians. Some of these have a great deal of cross-cultural validity. However, developmental norms paralleling physical growth norms are largely missing and are badly needed for assessment, screening, and detection of developmental risks. With better public health measures such as immunizations, significant increases have been achieved in the rates of infant and child survival globally. However, the circumstances that used to put infants at risk for death continue to put the survivors at risk for arrested or delayed development. Thus, the question of what happens to those who survive assumes great importance. With new policy shifts focusing on early childhood care and development, developmental psychologists are now being called on for their expertise. This is required especially for the cross-culturally valid conceptualization and assessment of human development.

Interactive conceptualization of nutrition, health, and psychosocial development, with each one acting on the other two, is proposed to better represent the dynamic processes of child development in the cultural context. For example, the effectiveness of health and nutrition programs can be increased by combining them with early childhood stimulation and education. The role of the proximal environment, particularly child care, is of importance here as a mediator of the macro-level influences. Thus, within the same poverty situations, some children thrive, whereas most others stay behind. Mother–child (or caretaker–child) interaction emerges as an important "shield" factor in promoting the child's resilience, conceptualized as "positive deviance." Here, supporting and training the mother or caretaker goes a long way toward leading to better child outcomes.

In dealing with disadvantaged contexts of child development, culturally valid assessment tools and developmental norms are crucially important. The challenge is to develop child development indicators that have cross-cultural validity to function as "standards" of development and also have cultural validity. Environmental indicators are also of crucial importance and are even less well established cross-culturally. For example, research points to the positive effects of maternal literacy

and education on child survival, longer intervals between pregnancies, better child physical growth, and more developmentally facilitative maternal child-rearing patterns. A concept such as "optimal parenting" becomes relevant here. When made applicable to large numbers of children and parents, such knowledge can be used for detection and screening of risk cases and for establishing and promoting the environmental factors conducive to healthy child development.

5.2. Schooling

Schooling is another case in point. A great deal of evidence points to the significant contributions of education, particularly schooling to societal development. Of greater relevance for psychology is the contribution of schooling for cognitive development. For example, Ceci showed that each additional month that a student remains in school may increase his or her IQ score above what would be expected if the student had dropped out. Other research provides similar evidence. Furthermore, a causal relationship is implicated from education to IQ, rather than the other way around, on the basis of higher-IQ children choosing to stay in school longer, as claimed by genotype–phenotype theory mentioned previously. There are different types of evidence supporting this causal relationship, including the effects of intermittent school attendance, delayed school entrance, length of school year, summer vacations, and early-year birth dates.

Given the overwhelming evidence regarding the societal and individual benefits of schooling, there needs to be strong advocacy for increased schooling for all, particularly for girls. Access to schooling is behind schedule in most of the world. For example, during the 1990s, when overall school enrollment reached 92% in North America, it remained at a mere 42% in Sub-Saharan Africa. Although it is developmental psychological research that reveals the positive effects of schooling on cognitive competence and human potential development, psychologists play only a marginal role in efforts to make "education for all" a reality in the world. At the same time, schools should be improved so as to entail better pedagogy and both culturally relevant and globally shared knowledge. Psychologists have a role to play here as well. For example, Stevenson and colleagues studied the cultural contexts of schooling and its cognitive products, focusing on the values and practices of Japanese, Chinese, and American schools and parents, with significantly different achievement outcomes for

children. The implications for educational policy and practice are far-reaching.

5.3. Early Enrichment

The importance of early experience for the development of competence is well recognized, notwithstanding the significance of organismic factors. Therefore, concerted efforts have been expended to provide deprived children with early enrichment in attempting to enhance their ability to benefit from formal schooling. "School readiness" is defined in terms of the child's activity level, social competence, psychological preparedness, and cognitive abilities, including preliteracy and prenumeracy skills. It is also reflected in the positive orientation and support of the child's family.

Research and applications during early childhood education abound in North America and Europe but are few and far between in developing countries. An important issue is whether a "center-based" or a "home/community-based" contextual approach is to be used (a related distinction is that between a child-focused and a parent-focused orientation). A general conclusion that can be drawn from research in various cultural contexts is that if the child's family environment can provide the child with the necessary stimulation, a center-based, child-focused approach is effective; otherwise, parental involvement is needed to improve the home environment. A home-based contextual approach has the potential to create lasting effects that would continue to support the child even after the completion of the early enrichment program, as evidenced by longitudinal research. It also tends to have multiple goals and multiple effects in supporting and changing the home environment, particularly the parents, together with the child. Given the crucial importance of the proximal environment and particularly of parenting, a contextual and culturally sensitive early enrichment can go a long way toward increasing the possibility that children from disadvantaged backgrounds can enter adulthood on a more equal footing with those from relatively advantaged backgrounds.

5.4. Problems in Development

Various problems in child development arise from environments that fall far short of presenting an optimal fit with children's developmental trajectories. Most of these emerge in social–cultural–economic deprivation (discussed earlier). Even more serious developmental problems may be encountered in especially difficult

circumstances such as war, family violence, child abuse and neglect, abject poverty involving malnutrition, debilitating and fatal morbidity (e.g., HIV/AIDS), and homelessness (e.g., street children). Under such difficult circumstances, beyond arrests in physical and cognitive/language development, emotional and social development also suffers greatly with possible psychopathology. Beyond medical and other research and interventions, there is a need for research that examines the cultural aspects involved such as societal values, parental views of children's needs, and parental behaviors and priorities for survival. Culturally sensitive assessment of the situation and interventions that are informed by cultural knowledge would be expected to have a greater chance for success. In particular, cultural strengths, such as closely knit human ties and social support mechanisms, could be identified and buttressed further, whereas vulnerabilities and weaknesses could be ameliorated. Thus, especially in the most difficult circumstances, a cultural/contextual approach is valuable. This is an important area of applied developmental psychology.

6. CONCLUSION

Child development occurs in culture and has to be studied in cultural context. Such a cultural/ecological perspective is necessary for both scientific and applied work. Understanding, assessing, and possibly enhancing child development requires that both a "whole child" perspective, attending to different spheres of development, and a contextual perspective be used. In so doing, both the unity and diversity of human development across cultures need to be recognized. Finally, developmental scientists should strive to go beyond understanding child development to use this understanding optimally in enhancing human development and well-being within and across cultures.

See Also the Following Articles

Diverse Cultures, Dealing with Children and Families from

Further Reading

Berry, J., Dasen, P. R., & Saraswathi, T. S. (Eds.). (1997). *Handbook of cross-cultural psychology: Basic processes and human development* (Vol. 2). Boston: Allyn & Bacon.

Bornstein, M. H. (Ed.). (1991). *Cultural approaches to parenting.* Hillsdale, NJ: Lawrence Erlbaum.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Bronfenbrenner, U., & Ceci, S. J. (1993). Heredity, environment, and the question "how": A first approximation. In R. Plomin, & G. E. McClearn (Eds.), *Nature–nurture and psychology* (pp. 313–324). Washington, DC: American Psychological Association.

Ceci, S. J., & Williams, M. W. (Eds.). (1999). *The nature–nurture debate: Essential readings.* Cambridge, MA: Blackwell.

Cole, M. (1996). *Cultural psychology: A once and future discipline.* Cambridge, MA: Belknap Press of Harvard University Press.

Gardiner, H., Mutter, J., & Kosmitzki, C. (1998). *Lives across cultures: Cross-cultural human development.* Boston: Allyn & Bacon.

Greenfield, P., & Cocking, R. (1994). *Cross-cultural roots of minority child development.* Hillsdale, NJ: Lawrence Erlbaum.

Kagitcibasi, C. (1996). *Family and human development across cultures: A view from the other side.* Hillsdale, NJ: Lawrence Erlbaum.

LeVine, R. A., Miller, P., & West, M. (1988). *Parental behavior in diverse societies.* (New Directions for Child Development, No. 40). San Francisco: Jossey–Bass.

Masten, A. S. (Ed.). (1999). *Cultural processes in child development.* Mahwah, NJ: Lawrence Erlbaum.

Nsamenang, B. (1992). *Human development in cultural context: A Third World perspective.* Newbury Park, CA: Sage.

Paris, S. C., & Wellman, H. (Eds.). (1998). *Global perspectives for education: Development, culture, and schooling.* Washington, DC: American Psychological Association.

Ryan, R. M., Deci, E. L., & Grolnick, W. S. (1995). Autonomy, relatedness, and the self: Their relation to development and psychopathology. In D. Cicchetti, & D. J. Cohen (Eds.), *Developmental psychopathology* (Vol. 1, pp. 618–655). New York: John Wiley.

Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, *9*, 545–570.

Child Testimony

Jodi A. Quas and Brianne E. Beck
University of California, Irvine, California, USA

GLOSSARY

direct questions Yes/No or short-answer questions that probe for particular details about an experienced or witnessed event; questions can range from nonleading to suggestive, depending on how the question is phrased and the type of information that is included.

free recall questions Vague general prompts requesting a narrative description of prior events; prompts provide minimal cues or information about what should be recalled.

script A general knowledge structure that represents information from a class of similar events; it is a description of the typical features of repeated events.

structured interviews Scripted interview prompts and questions that are designed to obtain as much narrative information from children and to be as minimally leading as possible.

suggestibility Susceptibility to errors when describing witnessed or experienced events; it can be caused by internal factors (e.g., biases) or external factors (e.g., questioning technique).

Since the 1980s, there has been a virtual explosion of studies concerning children's eyewitness abilities and participation in legal proceedings. The studies, along with considerable public, legal, and legislative attention, were motivated by a number of phenomena. One was a growing awareness of the widespread prevalence of maltreatment, especially child sexual abuse. Large U.S. national surveys revealed that many more individuals were reporting unwanted sexual experiences as children than had been recognized previously, yet few perpetrators were successfully prosecuted criminally. In response, increased efforts were made to identify victims of maltreatment and facilitate the prosecution of perpetrators of crimes against children. A second contributing factor involved revisions in statutes (e.g., competency requirements) that reduced barriers to admitting children's testimony as evidence in court. By reducing such barriers, the number of child victims/witnesses becoming involved in legal cases increased dramatically. Unfortunately, little was known about how to elicit accurate information (e.g., about abuse) from children. Legal professionals and scientists, as well as the public, focused on the consequences of failing to identify a victimized child rather than on the potential for false reports of a crime. A third phenomenon during the 1980s, however, altered the public's and researchers' focus on false allegations. Specifically, several high-profile legal cases that involved scores of children alleging sexual assault in their day care facilities took place. The children's allegations included fantastic and impossible claims and often emerged only after the children had been subjected to repeated, highly suggestive interviews. These children's reports, and the

cases in general, raised serious concerns about the reliability of children's memory abilities. Questions about how to protect children from harm and ensure that those who commit crimes against children are punished, while at the same time guarding against false allegations of abuse, were brought to the forefront of public, legal, and scientific attention. Research on children's memory, suggestibility, and false memories, as well as on children's participation in legal cases, has sought to answer these questions. The present article focuses on this research. First, research concerning factors that influence children's memory accuracy and susceptibility to false suggestions is presented. Second, the effects of participation in legal proceedings on children's well-being are discussed. Third, a brief overview of new directions in the field is provided. The article is limited in scope to findings that are most applicable to legal cases, are particularly well established, and have received considerable attention. It is also limited to discussion of child victims/witnesses. A large and growing body of literature now focuses on child defendants, and it is not possible to review both areas in a single brief article.

1. FACTORS THAT INFLUENCE CHILDREN'S TESTIMONY

Hundreds of studies have now been conducted concerning the reliability of children's eyewitness abilities. A number of factors have been identified that have implications for children's memory and suggestibility. Those discussed here can be heuristically divided into child, to-be-remembered event, and interview characteristics.

1.1. Child Characteristics

Numerous characteristics of children can influence their eyewitness abilities. The most well-studied characteristic is age, and the most consistent finding across research involves age differences in performance. With advancing age, children remember more information, are more accurate, are less suggestible, and make fewer false event reports. Both cognitive and social factors underlie age differences. Cognitively, older children are better able to conduct memory searches, know what is important to recount, and know how to cue their own memories. They also have more experiences and greater general knowledge, and this helps them to encode, interpret, and later remember information. Socially, older

children are more aware that adults can make mistakes, and this reduces their willingness to report false information suggested by adults as having occurred.

Other characteristics, in addition to age, have also been studied as predictors of children's memory and suggestibility. These include cognitive factors (e.g., intelligence, language, inhibitory control) and socioemotional factors (e.g., temperament, emotional reactivity, attachment). With regard to cognitive factors, although findings have not always been consistent, higher intelligence, greater language understanding, and better inhibitory control are generally associated with improved memory and reduced suggestibility, even when children's age is statistically controlled. With regard to socioemotional factors, outgoing temperament, low emotional reactivity, and secure attachment are often related to better memory, especially for stressful experiences.

Although research concerning child characteristics is important theoretically and provides considerable insight into processes underlying accuracies and inaccuracies in children's memory and suggestibility, it is difficult to apply to legal contexts. In particular, it is not possible to use any of the aforementioned characteristics, alone or in combination, as a litmus test to determine whether to accept or reject a particular child's eyewitness account. Nonetheless, the research may help to identify children who need special precautions taken when they are questioned about forensically relevant matters. Armed with information about factors that predict performance in individual children, interviewers may be better able to alter their questioning procedures to ensure that each child is interviewed in the most appropriate manner possible.

1.2. To-Be-Remembered Event Characteristics

Numerous factors associated with a to-be-remembered event may influence children's memory and suggestibility. Three factors that have received considerable attention are the stressfulness of the to-be-remembered event, the number of times that an event occurred, and children's level of involvement.

For some time, researchers have attempted to determine whether stress helps or hurts children's memory. Studies have focused on children's memory for mildly arousing laboratory-based events (e.g., fire alarms) as well as for highly distressing, salient, naturally occurring experiences (e.g., medical procedures, natural disasters). Although results have not revealed any

uniform positive or negative effect of stress on memory, some general conclusions are possible. First, children, like adults, remember emotional events better than they do neutral events. Thus, although the effects of varying levels of distress on memory remain unclear, stressful events appear to be more memorable than nonstressful neutral events. Second, memory for central details of an event (e.g., details directly associated with the cause of the stress) is often better than memory for peripheral details (e.g., information unrelated to the cause of the stress). It should be noted, however, that what is considered central to a child is not necessarily the same as what is considered central to an adult. For instance, a parent's departure during a medical procedure may be especially salient, and hence central, to a child even though the parent's departure is unrelated factually to the medical procedure. Third, the effects of stress on memory vary across children such that for some children (e.g., securely attached children) stress helps memory, whereas for other children (e.g., shy or insecurely attached children) stress inhibits memory. Researchers have begun to investigate interactions between stress and child characteristics as joint predictors of memory and suggestibility. The results of these new studies will continue to elucidate when and how stress affects children's performance.

The number of times an event is experienced can affect children's memory, although the direction of the effect varies depending on the type of to-be-remembered information. When an event is experienced repeatedly, children form scripts, which are mental representations of what "usually" happens or the common recurring features of the event. When asked to recount what happened during a repeated event, children can use these scripts to help them describe what usually happens. Thus, scripts enhance memory for the common repeated components. Unfortunately, children may fail to report nonrepeated details because these details are not included in the script of what usually occurs. Children may also have difficulty discriminating between repeated events and report details that are consistent with what usually occurs, even if those details did not occur during one particular incident.

A third event-relevant characteristic is children's participatory role. Children's memory is better for events in which they participated than for events that they merely witnessed. Active participation often results in more elaborate encoding and enhanced understanding of an event, both of which facilitate recall. Accordingly, children may often be better witnesses about events that they personally experienced than about events that they only witnessed, although some witnessed events, such as those that are especially relevant personally or are particularly salient or stressful, can also be remembered quite well.

1.3. Interview Characteristics

Much of the research concerning children's eyewitness testimony has focused on how characteristics of the interview affect performance. Results have been applied directly to legal contexts. For instance, several criminal convictions have been overturned because the interviews with alleged child victims were conducted in a manner that has been shown, in empirical research, to lead to errors and entirely false event reports. Also, interviewers are now often taught appropriate interviewing tactics and which tactics to avoid when questioning alleged child victims/witnesses.

How questions are phrased can affect the completeness and accuracy of children's reports. Questions are often divided into "free recall" and "direct" questions. Free recall questions are vague general statements prompting a child to recount, in narrative form, a past event (e.g., "Tell me what happened when . . ."). Direct questions include nonleading, leading, and suggestive questions. Nonleading questions do not imply incorrect information or a particular response (e.g., "What color was the car?"). Leading questions imply a particular response (e.g., "Was the car red?"). Suggestive questions contain false embedded information (e.g., "What was the man wearing when he entered the room?" when in fact there was no man) or explicitly imply which answer should be given ("He entered the room first, didn't he?" when the man did not in fact enter first).

Free recall questions are less controversial than direct questions because children's free recall responses are typically quite accurate. That is, children rarely provide incorrect information when asked, in an open-ended manner, to recount a particular event. The major drawback to free recall, however, is that children's reports are often incomplete. Children might not know what is important to recount. They also might have difficulty in forming cogent narrative answers, even when they remember event details.

To obtain additional information, direct questions that probe for specific information are often needed. Such questions indeed increase the amount of information that children provide; however, the questions, especially those that are suggestive, also increase errors. Children may answer in a way that they believe the interviewer desires, they might not be comfortable

indicating that they do not know an answer, they may develop a "yeah"- or "nay"-saying response bias, or they may accept the interviewer's suggestions as fact. Each of these increases errors, and in some circumstances it may lead children to claim that entirely false events occurred.

Importantly, although children's free recall reports are generally accurate but incomplete and children's responses to direct questions are less accurate, two exceptions to these trends must be mentioned. First, accuracy of free recall can be compromised if, prior to asking free recall questions, children are exposed to false suggestive information. For instance, before an interview, a parent or an interviewer might imply to a child that an alleged perpetrator did something bad or a child might learn from other children what happened during an alleged event. These preinterview sources of false suggestions lead to increased errors in children's responses to both free recall and direct questions. Second, when no abuse occurred, children tend to be more accurate when answering abuse-related direct questions (e.g., "Did he take pictures of you when you were in the bathroom?," "Did he touch your bottom?" [asked to children who merely experienced a play interaction with an adult]) than when answering non-abuse-related questions. Thus, children's willingness to err in response to direct questions varies depending on the content of those questions.

Highly leading interview instructions can have a dramatic effect on children's memory and suggestibility. Such instructions include asking children to imagine or pretend that a particular event occurred or introducing false stereotypic information about an event or individual(s). When these instructions are given before or during an interview, children will often alter their statements about an individual's activities or change their interpretation of an event and recount information that is consistent with the leading instructions. In addition, when children are explicitly told that an event occurred and are given information about what happened, a sizable minority will repeat the information as having happened even if the event is false. Interestingly, children are less likely to assent that false negative events (e.g., falling off a tricycle and getting stitches) occurred than to assent that false positive events (e.g., going on a hot air balloon ride) occurred, although some children nonetheless assent to both positive and negative events. What remains of theoretical interest is the extent to which children believe that the false events took place versus the extent to which they are simply acquiescing to an interviewer's suggestions.

Question repetition, both within and across interviews, has been especially controversial. Some claim that simply repeating questions about untrue information causes false reports in children. However, such a phenomenon has not yet been confirmed empirically. Instead, the effects of repeated questions vary as a function of the type of question, whether questions are repeated within or across interviews, and preinterview instructions. Within an interview, particularly young (e.g., preschool-age) children often change their responses when asked repeated direct (e.g., yes/no) questions. They tend to interpret the repetition as an indication that their first response was incorrect. In contrast, when questions are repeated across interviews, at least when no highly leading instructions have been provided, earlier interviews can actually have a buffering effect, reminding children of an original event and reducing errors in later interviews. Finally, when repeated interviews are combined with highly leading interview instructions such as those described previously, a high percentage of children (often one-third or more) will make false reports that are consistent with the instructions.

1.4. Summary

A variety of factors have been studied as predictors of accuracies and inaccuracies in children's memory and eyewitness abilities. Some findings (e.g., those concerning age differences in performance, the accuracy of children's free recall reports, the deleterious effects of highly suggestive interview questions and instructions) have been consistently found in empirical research and are now met with minimal controversy when applied to legal settings. Other findings (e.g., those involving effects of stress on memory or interview repetition), however, have been more variable and are not amenable to broad straightforward conclusions or applications to legal contexts. That is, whether many factors (e.g., repeated interviews, stress, participation) help or hurt children's eyewitness abilities depends on a number of considerations, such as how questions are asked and the type of memory being probed. These considerations need to be taken into account when attempting to evaluate child victims'/witnesses' reports.

2. CHILDREN'S INVOLVEMENT IN THE LEGAL SYSTEM

Although most research concerning children's testimony has focused on the accuracy of children's eyewitness

abilities, studies also have investigated the effects of legal involvement on children and the use of innovative practices to facilitate children's participation in legal cases. Some psychologists and legal scholars contend that children suffer "secondary" victimization by participating in a legal system that was not designed for, and can be quite intimidating to, children. Such individuals advocate altering standard courtroom procedures (e.g., allowing children to testify without a defendant present) to reduce children's distress. Other scholars, however, maintain that, because all witnesses can be distressed during a legal case, distress should not be used as the basis for abandoning defendants' fundamental rights or changing standard legal practice.

A few studies have investigated the effects of legal involvement on children. Most noteworthy are studies of the effects on child sexual abuse victims. Findings indicate that some children are adversely affected by legal involvement in terms of increased distress and behavior problems following their legal cases. Factors that increase a child's risk of adverse effects include testifying multiple times, testifying while facing the defendant, not having maternal support during the case, and not having corroborative evidence confirming the abuse allegation. However, not all children appear to be adversely affected, and adverse effects appear more often following criminal involvement than following juvenile or family court involvement. Furthermore, the extent to which adverse effects are maintained over long time periods has only recently begun to be investigated.

As mentioned previously, some advocate altering standard legal procedures to facilitate children's participation in legal cases. A related line of research has investigated the use of these alterations and their effects on children's distress and case outcomes. Innovative techniques include those that affect pre-court interviews with children (e.g., multidisciplinary centers, videotaping interviews) and those that affect in-court practices (e.g., admitting children's out-of-court statements [hearsay] as evidence, allowing children to testify via closed-circuit television testimony). Although there is support for the use of the innovations, and the constitutionality of their use has been upheld by some U.S. courts, their actual use appears to be limited. In particular, surveys of prosecutor and judicial behaviors indicate that techniques that alter in-court practices are rarely used, primarily due to concerns that decisions will be overturned. Techniques that alter how children are interviewed are also rated as too costly for widespread use. Despite the limited use of innovations, however, they appear to be effective

in reducing children's distress. Also, studies of jurors' decision making have revealed relatively few differences in case outcomes based on whether children testify live or children's statements are presented in some other format (e.g., hearsay is presented, children testify via closed-circuit television). Finally, the use of several innovations, such as videotaped testimony and closed-circuit television testimony, is standard in other countries (e.g., Great Britain). Perhaps as awareness of both the benefits of the innovations and the lack of adverse effects on case outcomes grows, the innovations will be used more frequently in the United States as they are in other countries.

3. CONCLUSIONS

Numerous characteristics, including those both intrinsic (e.g., age) and extrinsic (e.g., stress of a to-be-remembered event, interviewer demeanor, testifying in court) to children, have implications for children's ability to recount past experiences accurately and for their ability to cope with involvement in the legal system. Although this article has highlighted some of the better established and particularly noteworthy findings and controversies, it is by no means exhaustive. Furthermore, several exciting new directions in research continue to refine understanding of the conditions under which children's testimony is more likely to be accurate and how to facilitate children's participation in legal cases. These new directions focus on how characteristics of children interact with contextual characteristics to affect their memory and suggestibility and on the development and testing of formal interview protocols that can be used in forensic contexts to ensure that children are questioned in the most appropriate manner. With regard to interactions between children and contextual characteristics, it is likely that individual children respond very differently to particular events, such as stressful experiences, and these different responses may affect how individual children later remember the events. With regard to interview training protocols, simply providing forensic interviewers with background knowledge (e.g., about the hazards of leading questions) and general training does not improve how actual interviews are conducted. Thus, researchers have begun to develop specific guidelines and structured protocols that interviewers can use when questioning alleged child victims/witnesses. Although these new protocols, which are more widely accepted in other countries (e.g., Great Britain, Israel), are currently being tested in the United

States, they already appear to be effective in increasing the amount of narrative information provided (e.g., in response to free recall questions) and in helping interviewers to avoid the most problematic questioning tactics. These protocols may also decrease the number of times children need to be questioned, thereby reducing distress that can result from being interviewed multiple times about alleged crimes.

As new psychological research is conducted, and as previously reported findings are confirmed and extended, it will be imperative to continue to educate those involved in actual forensic contexts regarding abilities and limitations of child witnesses. Not only will this knowledge advance scientific understanding of children's memory capabilities and suggestibility, but it will also benefit both alleged victims/witnesses and defendants in legal cases involving children.

See Also the Following Articles

Eyewitness Testimony ▪ Interrogation and Interviewing

Further Reading

Ceci, S. J., & Bruck, M. (1998). Children's testimony: Applied and basic issues. In W. Damon, I. E. Sigel, & K. A. Renninger (Eds.), *Handbook of child psychology* (5th ed., Vol. 4, pp. 713–773). New York: John Wiley.

Ceci, S. J., & Friedman, R. D. (2000). The suggestibility of children: Scientific research and legal implications. *Cornell Law Review*, 86, 33–109.

Goodman, G. S., Emery, R., & Haugaard, J. (1998). Developmental psychology and the law: Divorce, child maltreatment, foster care, and adoption. In I. Sigel & A. Renninger (Eds.), *Handbook of child psychology*, Vol. 4: *Child psychology in practice* (5th ed.). New York: John Wiley.

Goodman, G. S., Taub, E., Jones, D. P. H., England, P., Port, L. P., Rudy, L., & Prado, L. (1992). Emotional effects of criminal court testimony on child sexual assault victims. *Monographs of the Society for Research in Child Development,* 57 (Serial No. 229).

Home Office Communication Directorate. (2001). *Achieving best evidence in criminal proceedings: Guidance for vulnerable or intimidated witnesses, including children*. London: Author.

Lyon, T. D. (1999). The new wave of suggestibility research: A critique. *Cornell Law Review*, 84, 1004–1087.

Pool, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.

Quas, J. A., Goodman, G. S., Ghetti, S., & Redlich, A. (2000). Questioning the child witness: What can we conclude from the research thus far? *Trauma, Abuse, and Violence*, 1, 223–249.

Classical Conditioning

P. Ferrándiz

Catedrática, Universidad Complutense, Madrid, Spain

GLOSSARY

behaviorism The term originated with the work of Watson. Behaviorism was a movement in psychology that argued that mental events were impossible to observe and measure. It was claimed that psychology was only the study of observable behavior. Until the 1960s, this was the dominant approach to learning.

cognitivism An approach in psychology that was born in 1956. It began as a reaction to behaviorism. It proposed that human information and storing processes were very important in the process of learning.

conditioned or conditional response (CR) Any response that is learned by conditioning.

conditioned or conditional stimulus (CS) Any stimulus that, through conditioning, ultimately evokes a conditioned response. The conditioned response is originally a neutral stimulus.

contiguity The quality of two events occurring relatively closely in time. Conditioning will only occur if events or stimuli (CS–US) follow each other.

contingency The relationship between two events in which one depends on the other. Contingency refers to the probability that when the CS occurs, the US is likely to follow.

unconditioned or unconditional response (UR) Any response elicited spontaneously by an unconditioned stimulus, without any prior learning or conditioning.

unconditioned or unconditional stimulus (US) Any stimulus that elicits a response without any previous learning.

The simplest associative mechanism, whereby organisms learn to produce new responses to stimuli and learn about relations between stimuli or events, is classical, or Pavlovian, conditioning. In classical conditioning, a neutral stimulus is repeatedly paired with another stimulus, called an unconditioned stimulus (US), that naturally elicits a certain response, called an unconditioned response (UR). After repeated trials, the neutral stimulus becomes a conditioned stimulus (CS) and evokes the same or a similar response, now called the conditioned response (CR). Most of the concepts presented in this article were developed by Pavlov. However, it also examines how original classical conditioning has been modified. Although the basic conditioning paradigm emerged 100 years ago, its comprehension and conceptualization are still developing.

1. FROM PAVLOV TO THE PRESENT

Russian physiologist I. Pavlov (1849–1936) was one of the pioneers in research on classical conditioning. His investigations, carried out at the end of the 19th

century, form the axis of associative learning. Pavlov was interested in the digestive system and, through these studies, he observed that dogs, the subjects of his experiments, anticipated the salivation response when they saw food. Subsequently, Pavlov presented a light or the ticking of a metronome [conditioned stimulus (CS)] for a number of seconds before the delivery of food [unconditioned stimulus (US)]. At first, the animal would show little reaction to the light but, as conditioning progressed, the dog salivated during the CS even when no food was delivered. This response was defined as the conditioned response. Hence, research on classical conditioning began. Soviet psychology has changed in many ways since the death of Pavlov in 1936, and many of these changes brought it closer to American psychology. One way in which the general theory of Pavlovian conditioning developed similarly in the two countries was via the incorporation of cognitive variables. The originally simple Pavlovian paradigm was expanded and reinterpreted through the efforts of many American researchers. American learning theorists developed alternative methods for estimating the strength of the association between a CS and a US. They also identified new Pavlovian phenomena and suggested new explanations of the basic mechanisms underlying the conditioning produced by use of Pavlovian procedures. Although Pavlov's writings were widely disseminated in the United States, most of the criticism of his ideas was expressed in 1950 by J. A. Konorski in Poland. Konorski and Miller, Polish physiologists, began the first cognitive analysis of classical conditioning, the forerunner of the work by R. A. Rescorla and A. R. Wagner (American psychologists) and A. Dickinson and H. J. Mackintosh (English psychologists).

In the 1950s, some clinical psychologists—such as J. Wolpe (in South Africa), H. J. Eysenck (in the United Kingdom), and B. F. Skinner (in the United States)—began to treat psychological disorders with a technology based on learning principles, aiming to substitute pathological behavior for a normal one. This movement, known as behavior modification, was based on both classical and operant conditioning and reinforced the popularity of these techniques among clinicians.

Significant changes on theoretical views upon this phenomenon were introduced by Rescorla and Wagner in 1972, whose model assumed that organisms, in their interaction with the environment, form causal expectations that allow them to predict relations between events. From this perspective, learning would consist of the acquisition of information about the causal organization of the environment. In order to acquire this information, associations between elements are established. However, it is not sufficient for the elements to be contiguous so that an animal will associate them, as Pavlov proposed; it is also necessary for the elements to provide information about a causal relation. In fact, the central concept is the notion of contingency and the way in which it is represented in the animal's mind. This new "cognitive associationism" accepted the existence of both cognitive processes and associative mechanisms in the mind and rapidly gained wide support among learning experimentalists.

Although in some cases the organism seems to be conditioned only on grounds of CS–US "contiguity," in many other cases learning seems to allow the organism to predict US arrival from the CS appearance; this would be due to a "contingency relationship" that is expressed in terms of probability; if the probability of the appearance of the US is the same in the presence and in the absence of the CS, contingency will be null and no conditioning will occur. On the contrary, if the probability of the US appearing in the presence of the CS is greater than the probability of it appearing without the CS, then conditioning occurs. Therefore, classical conditioning is defined as the capacity of organisms to detect and learn about the predictive relation between signals and important events. Currently, classical conditioning is not considered to be simple and automatic learning about the establishment of CS–US relations but rather a complex process in which numerous responses intervene, not just glandular and visceral responses, as was thought at first. It also depends on the relevance of the CS and the US, the presence of other stimuli during conditioning, how surprising the US may be, and other variables.

Lastly, the growth of connectionist modeling has renewed interest in classical conditioning as a fruitful method for studying associative learning.

2. BASIC CONDITIONING PROCEDURE

The procedure of classical conditioning consists of the repeated presentation of two stimuli in temporal contiguity. First, a neutral stimulus (NS) is presented— that is, a stimulus that does not elicit regular responses or responses similar to the unconditioned response (UR). Immediately after that, the US is presented.

NS (a bell) → OR (orientation response)

US (presence of food in the dog's mouth) → UR (salivary response)

CS (a bell) → CR (salivary response)

FIGURE 1   Diagram of the sequence of events in classical conditioning.

Because of this pairing, the NS will become a CS and, therefore, will be capable of provoking a conditioned response (CR) similar to the UR that, initially, only the US could elicit (Fig. 1).

On the initial trials, only the US will elicit the salivation response. However, as the conditioning trials continue, the dog will begin to salivate as soon as the CS is presented. In salivary conditioning, the CR and the UR are both salivation. However, in many other conditioning situations, the CR is very different from the UR. According to Pavlov, the animals learn the connection between stimulus and response (CS–UR). Currently, it is understood that animals learn the connection between stimuli.

3. BASIC PROCESSES

Some basic processes that affect all sorts of conditioning learning have been identified. They include the following:

1. Acquisition: the gradual increase in strength of CR, linking the two stimuli (CS followed by the US).

2. Extinction: the reduction and eventual disappearance of the CR at the CS onset, after repeated presentations of CS without being followed by US.

3. Generalization: when a CR is linked to a CS, the same CR is elicited by other CSs in proportion to their similarity to the original CS.

4. Discrimination: occurs when subjects have been trained to respond in the presence of one stimulus but not in the presence of another.

4. EXPERIMENTAL PROCEDURES

4.1. Appetitive Conditioning

Appetitive conditioning is a modality in which the US (and CS) will originate a positive (approaching, consummative) behavior in the organism.

4.2. Aversive Conditioning

In aversive conditioning, the US and CS have negative properties and will generate an escape or evitation behavior. Aversion to some taste (Garcia effect) showed the possibility of obtaining long aversive reactions to some CS taste that, after being linked to some US with strong aversive digestive qualities, would generate an evitation response after a long interstimuli interval. Learning taste aversion plays an essential role in the selection of food because it prevents the ingestion of toxic foods. It has also helped to understand and treat food aversions in people suffering from cancer and undergoing chemotherapy.

5. EXCITATORY CONDITIONING

Excitatory conditioning occurs because the CS allows the subject to predict when the US will occur. In excitatory conditioning, there is a positive relation between the CS and the US. The excitatory stimulus (CS+) becomes a stimulus that elicits a CR after its association with the US. Traditionally, classical conditioning research has focused on this kind of conditioning.

6. INHIBITORY CONDITIONING

Inhibitory conditioning occurs because the CS allows the subject to predict that no US will be presented; that is, an inhibitory conditioned stimulus (CS-) signals the absence of an unconditioned stimulus (either positive or aversive), provoking the inhibition of the CR in the organism. This kind of conditioning was studied by Pavlov, and there has been a renewed interest that continues today.

One of the clearest examples of this procedure is the one presented by M. Domjan: A red traffic light (CS+) signals the potential danger (US), which is the car. If a policeman indicates that we can cross the street (CS-) despite the red light, he is signaling that there is no danger (US). Inhibitory stimuli, such as "out of order" or "do not cross" signs, provide useful information.

7. WHAT INFLUENCES CLASSICAL CONDITIONING?

The capacity of stimuli to become associated is modified by many phenomena, such as the CS–US interval, spatial

contiguity between stimuli, the similarity of the CS and the US, the intertrial interval, the characteristics of the CS and the US (intensity, novelty, and duration), or previous experience of the CS (latent inhibition), the simultaneous presence of other more salient stimuli (overshadowing), or more informative stimuli (blocking).

7.1. CS–US Interval

The interstimulus interval is the time between the appearance of the CS and the appearance of the US. It is one of the critical factors that determine the development of conditioning. Different CS–US interstimulus intervals are usually employed.

Simultaneous conditioning describes the situation in which the CS and the US begin at the same time. Delayed conditioning occurs when the onset of the CS precedes the US onset. Trace conditioning occurs when the CS and US are separated by some interval in which neither of them is present. The US is not presented until after the CS has ended. Backward conditioning occurs when the CS is presented after the US. This procedure casts doubt on the principle of stimulus contiguity because although the CS and the US are presented in inverse order, the association should be equally efficient.

The simultaneous and backward conditioning procedures produce poorer conditioning in comparison to delayed and trace procedures. However, today, due to the use of more sophisticated measurements than those traditionally employed, all the procedures can produce efficient conditioning.

8. APPLICATIONS

Classical conditioning is relevant to behavior outside the laboratory in daily life and plays a key role in understanding the origins and treatment of clinical disorders. In humans, classical conditioning can account for such complex phenomena as an individual's emotional reaction to a particular song or perfume based on a past experience with which it is associated; the song or perfume is a CS that elicits a pleasant emotional response because it was associated with a friend in the past. Classical conditioning is also involved in many different types of fears or phobias, which can occur through generalization.

Pioneering work on this type of phenomena was carried out by J. B. Watson and collaborators in the 1920s. They were able to condition a generalized phobia to animals and a white furry object in a young child, "Albert" (creating a learned neurotic reaction), and also to extinguish a phobia in another child, "Peter," through a learning procedure (carried out by Mary C. Jones). Both cases have been considered pioneering work in the behavior modification field.

8.1. Everyday Human Behavior

8.1.1. Classical Conditioning and Emotional Response

Recent reanalyses have evidenced the relevance of individual differences and cognitive variables. People who are more anxious in general develop phobias from experiences that do not produce fear in others.

Fears can be learned by multiple means that include direct Pavlovian conditioning, but the Pavlovian paradigm, with the revisions of the neoconditioning theory, continues to be one of the most influential theories on phobic anxiety. All fears or phobias do not seem to be explainable by simple Pavlovian conditioning, although many may be. Many phobias are reported to be the result of indirect or cognitive learning, including learning by observation and by information transfer. Classical conditioning has also been useful to explain panic attacks.

8.1.2. Classical Conditioning and Food Aversion

Food aversions are a serious clinical problem, especially those developed by association with the nausea produced by chemotherapy for cancer, which may later interfere with eating many foods. Applying modern knowledge of the classical conditioning process, treatment usually consists of trying to prevent the aversion by minimizing the predictability relationship of the CS and the US (nausea). Giving the patients a food with a highly salient, novel taste in novel surroundings results in blocking, and the novel cues overshadow the tastes of more familiar foods. The result is a strong aversion to cues that will not be present after chemotherapy is finished. This is helpful and causes no problems because there is little development of conditioned nausea to familiar foods when eaten at home.

8.2. Clinical Treatment

Behavior therapy, based on the principles of classical conditioning, has been used to eliminate or replace behavior, to eliminate the emotional responses of fear

and anxiety, and as treatment for nocturnal enuresis, alcoholism, and so on.

8.2.1. Systematic Desensitization Therapy

As behavioral methods developed over time, a behavior therapy technique called systematic desensitization was devised based broadly on the classical conditioning model. Undesirable responses, such as phobic fear reactions, can be counterconditioned by the systematic desensitization technique. This technique inhibits expressions of fear by encouraging clients to face the feared CS and thus allowing extinction to occur. In systematic desensitization, anxiety is associated with a positive response, usually relaxation. Systematic desensitization is a procedure in which the patient is gradually exposed to the phobic object; training in progressive relaxation is an effective and efficient treatment for phobias.

8.2.2. Implosive (Flooding) Therapy

One approach to treating phobias with classical conditioning was originally called implosive therapy (flooding). It is used to extinguish the conditioned fear response by presenting the CS alone, repeatedly, and intensely. The phobic individual experiences the CS, and all the conditioned fear is elicited, but no aversive US follows, nothing bad happens to the subject in the presence of the CS, and so the conditioned fear of the CS disappears.

See Also the Following Articles

Cognitivism

Further Reading

Chance, P. (2002). *Learning and behavior* (5th ed.). New York: Brooks/Cole.

Davey, G. (Ed.) (1987). *Cognitive processes and Pavlovian conditioning in humans*. Chichester, UK: Wiley.

Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge, UK: Cambridge University Press.

Domjan, M. (2000). *Essentials of conditioning and learning* (2nd ed.). Belmont, CA: Wadsworth/Thomson.

Domjan, M. (2003). *Principles of learning and behavior* (5th ed.). Belmont, CA: Wadsworth/Thomson.

Klein, S. B. (2001). *Learning: Principles and applications* (4th ed.). New York: McGraw-Hill.

Leahey, T. H., & Harris, R. J. (2000). *Learning and cognition* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Leslie, A. M. (2001). Learning: Association or computation? *Current Directions in Psychological Science*, *10*(4), 124–127.

Lieberman, D. A. (2000). *Learning: behavior and cognition* (3rd ed.). Pacific Grove: CA: Wadsworth.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford, UK: Clarendon.

Mackintosh, N. J. (Ed.). (1994). *Animal learning and cognition*. New York: Academic Press.

Mazur, J. E. (2002). *Learning and behavior* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

O'Donohue, W. (1998). *Learning and behaviour therapy*. Boston: Allyn & Bacon.

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.

Schwartz, B., Wasserman, E. A., & Robbins, S. J. (2002). *Psychology of learning and behavior* (5th ed.). New York: Norton.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning: II. Current research and theory*. New York: Appleton-Century-Crofts.

Clinical Assessment

Irving B. Weiner

University of South Florida, Tampa, Florida, USA

Clinical assessment is a procedure in which various forms of information drawn from interviews, behavioral observations, standardized tests, collateral reports, and historical documents are integrated to facilitate decisions about diagnosis and treatment planning.

GLOSSARY

deception A conscious and deliberate attempt to appear less disturbed and more capable than is actually the case; it is also known as "faking good."

idiographic A way of describing people according to their individual and unique characteristics.

malingering A conscious and deliberate attempt to appear more disturbed and less capable than is actually the case; it is also known as "faking bad."

nomothetic A way of categorizing people according to characteristics they share in common with other people.

performance-based measures Assessment instruments that sample how people respond when given a task to do.

self-report inventories Assessment instruments that sample how people describe themselves when asked to indicate whether, or to what extent, various statements apply to them.

standardized tests Formal measuring instruments that consist of a specific set of stimuli and instructions and that are always presented to respondents in the same way.

structured interviews Interviews based primarily on a preset schedule of specific questions.

unstructured interviews Interviews consisting primarily of open-ended inquiries and comments.

1. NATURE AND PURPOSE OF CLINICAL ASSESSMENT

Clinical assessment consists of collecting various kinds of information about the nature of people to facilitate arriving at certain kinds of decisions intended to promote the welfare of persons being assessed. These decisions most commonly address interrelated issues of differential diagnosis and treatment planning, and they typically are reached in consultation with persons who have been assessed concerning what would be in their best interest.

1.1. Differential Diagnosis

Treatment of persons with physical or psychological problems proceeds best when the nature and origins of these problems can be identified with reasonable certainty. Adequate differential diagnosis includes both nomothetic and idiographic components. The nomothetic component consists of categorizing the type of disorder a person has and noting the kinds of characteristics the person shares in common with other people who have the same disorder. The idiographic

component consists of specifying the particular life experiences that have given rise to the disorder in this individual case and identifying current circumstances that appear to be perpetuating the disorder.

1.2. Treatment Planning

Differential diagnostic information derived from clinical assessments guides treatment planning in four key respects. First, the category of disorder with which a person is diagnosed calls attention to modes of treatment that have been shown to be effective in treating people with the same or a similar category of disorder. Second, the diagnosed severity of a person's disorder indicates whether treatment can be provided adequately on an outpatient basis or instead requires a hospital or other residential setting. Third, diagnostic information about past and present events in a person's life often points to topics and goals on which the treatment sessions should focus. Fourth, in addition to identifying critical treatment topics and goals, diagnostic information can alert therapists to patient characteristics that might slow progress in the treatment, at least until these potential obstacles to progress have been minimized or overcome.

2. SOURCES OF INFORMATION IN CLINICAL ASSESSMENT

Clinical assessment is a multifaceted process in which information is drawn from diverse sources. These sources of information include interviews, behavioral observations, standardized tests, collateral reports, and historical documents. The data obtained from these sources sometimes converge, and in such cases the confirmatory information they provide broadens the base of the assessment process and lends certainty to inferences drawn from it.

In other instances, diverse sources of information may diverge, for example, when a man being interviewed describes himself as a laid back person who is easy to live with, whereas his wife (a collateral informant) describes him as an irritable and demanding person who is nearly impossible to live with. Instances of divergent data from different sources often help clinical assessors to sharpen their decision making by identifying why these divergences have occurred and determining which sources have yielded the most reliable and useful information. Ordinarily, prior to collecting assessment data, it is difficult to anticipate which sources of information will converge (and thus be redundant) and

which will diverge (and thus deepen understanding of the person being assessed).

2.1. Interviews

Mental health specialists who conduct clinical assessment interviews ask patients questions and invite them to talk about their symptoms and problems, any previous treatment they have received, their family backgrounds, and their developmental, educational, occupational, and social histories. Depending on the preferences of the clinician and the capacities of the respondent, these interviews may be relatively structured or relatively unstructured. Relatively structured assessment interviews are based primarily on a preset schedule of specific questions (e.g., "How many brothers and sisters do you have?," "Do you have trouble sleeping at night?"). Relatively unstructured interviews, in contrast, consist primarily of open-ended inquiries and comments (e.g., "What was growing up like for you?," "You seem a little anxious right now"). In actual practice, most clinical assessment interviewers combine some open-ended inquiries with various specific questions intended to clarify items of information.

2.2. Behavioral Observations

Clinical assessors learn a great deal about people from observing how they sit, stand, walk, gesture, show their feelings, and express themselves in words. These observations are made mostly while conducting assessment interviews, but clinicians sometimes have opportunities to observe people in other contexts as well, for example, as patients in a residential treatment setting or as inmates in a correctional facility.

Regardless of their context, behavioral observations can take a variety of forms. For example, clinicians working with families may take note of how family members talk to each other during a family group interview or how they interact during a home visit, each of which is a sample of fairly natural behavior. Alternatively, clinicians may watch and listen to how family members collaborate when they are given some task to do or are asked to decide on a single set of answers to a questionnaire, each of which is an artificially contrived situation.

2.3. Standardized Tests

Standardized tests are formal measuring instruments that consist of a specific set of stimuli and instructions and that are always presented to respondents in the same

way. The systematized materials and procedures that characterize standardized tests make it possible to collect normative reference data concerning how certain groups of people (e.g., men, women, children, Americans, Japanese) are likely to respond to these tests on average.

The standardized tests commonly used in clinical assessments are categorized as being either self-report inventories or performance-based measures. Self-report inventories sample how people describe themselves when asked to indicate whether, or to what extent, various statements apply to them (e.g., "I worry quite a bit over possible misfortunes," "In school, I was sometimes sent to the principal for bad behavior"). Performance-based measures sample how people respond when given a task to do (e.g., putting blocks together to make various designs, describing what a series of inkblots look like). Conjoint use of self-report inventories (what people say about themselves) and performance-based measures (how people go about doing assigned tasks) typically maximizes the scope and utility of the assessment data obtained with standardized testing.

2.4. Collateral Reports

In addition to interviewing, observing, and testing people being evaluated, clinical assessors may obtain collateral information from other individuals who know them or have observed them. Collateral informants are most commonly family members, but they may also include friends, teachers, and employers. Collateral reports usually enhance the scope and reliability of clinical assessments, especially when the people being assessed are unable or reluctant to talk freely or to cooperate fully with testing procedures.

At times, however, mistaken impressions on the part of collateral informants may detract from the accuracy of their reports, which turn out to be less dependable than what the people being assessed have to say. Informants' reports may also be of questionable reliability if personal issues they have with the people being assessed bias their perspectives. Clinical assessors may also confront instances in which people being evaluated exercise their rights to confidentiality by denying access to informants other than themselves or to historical documents.

2.5. Historical Documents

Historical documents provide real accounts of events in people's lives for which their own recollections may be vague or distorted. School, medical, military, and criminal records (when available) enhance the dependability of clinical assessments by documenting life experiences that people may be reluctant to talk about or unable to remember clearly. In addition, when there is concern that people being assessed have suffered some decline in mental or emotional capacity from previously higher levels of functioning, historical documents can provide a valuable baseline for comparison. This is especially the case when the past records include clinical assessments.

3. ISSUES IN CLINICAL ASSESSMENT

Contemporary clinical assessments must typically take account of four issues, especially with respect to collecting and using standardized test data. First, are the test results reliable and valid? Second, could the test results reflect malingering or deception on the part of the respondent? Third, are the test findings applicable in light of the respondent's cultural context? Fourth, can one make appropriate use of computer-based test interpretations?

3.1. Reliability and Validity

The reliability and validity of test results refer to whether the results are dependable, that is, whether they provide an accurate measurement of whatever they are measuring. The reliability of tests is most commonly estimated by examining the internal consistency of their items or by determining the retest stability of the scores they generate. The validity of tests refers to how much is known about what they do in fact measure, as determined by the extent to which their scores correlate with other phenomena such as a diagnosed condition and a favorable response to a particular form of treatment.

Currently available clinical assessment instruments vary in their demonstrated reliability and validity, and examiners must be cautious about relying on measures that are limited in these respects. On balance, however, research findings indicate that most of the tests commonly used in clinical assessment, when administered and interpreted properly, meet the criteria for acceptable reliability and validity.

3.2. Malingering and Deception

Malingering among persons being assessed consists of a conscious and deliberate attempt to appear more

disturbed and less capable than is actually the case. Malingering, also known as "faking bad," commonly emerges as an effort to minimize being held responsible for some misconduct (e.g., being found not guilty of a crime) or to maximize being held deserving of receiving some type of recompense (e.g., being awarded damages for an alleged injury).

Deception, in contrast, consists of attempting to appear less disturbed and more capable than is actually the case. Deception, also known as "faking good," usually arises as an attempt to gain some benefit (e.g., being employed in a position of responsibility) or to avoid losing some privilege (e.g., being deprived of child custody or visitation rights).

Clinical assessors can draw on various guidelines for identifying malingering and deception on the basis of interview behavior and responses to standardized tests. Based on criteria related to consistency versus inconsistency, these guidelines include specific kinds of interview questions, validity scales on standardized tests, and certain standardized tests designed specifically to reveal efforts to fake bad or fake good. Used in conjunction with data from observational, collateral, and historical sources, these interview and test measures are usually effective in detecting malingering and deception. Nevertheless, research findings indicate that resourceful and motivated respondents, especially when they have been coached with respect to what they should say and how they should conduct themselves, can succeed in misleading even experienced examiners.

3.3. Cross-Cultural Contexts

Cross-cultural clinical assessment generally proceeds in two phases. The first phase consists of drawing on sources of assessment information to identify the nature of an individual's psychological characteristics. The second phase consists of determining the implications of these psychological characteristics within the person's cultural context. With respect to collecting information, some assessment methods are relatively culture free (e.g., copying a design from memory). Other assessment methods, especially those involving language (e.g., interviews, self-report inventories), are relatively culture bound. To obtain reliable results, particularly with culture-bound methods, assessment data should be collected by an examiner or with a test form in a language with which the person being examined is thoroughly familiar.

Following collection of the data in cross-cultural assessment, the process of interpreting these data should be attuned as much as possible to available normative standards within groups to which the person being examined belongs. On the other hand, normative cross-cultural differences in assessment findings, such as scores on some test variable, may indicate an actual difference between cultures in the psychological characteristic measured by that variable rather than a need to adjust standards for interpreting the variable. In the case of standardized tests, in fact, many variables identify psychological characteristics in the same way regardless of the person's cultural context.

The main challenge in doing adequate cross-cultural assessment arises during the second phase of the process, that is, when clinicians must estimate the implications of a person's identified psychological characteristics for how the person is likely to function in his or her daily life. In some instances, psychological characteristics are likely to have universal implications. For example, being out of touch with reality will ordinarily interfere with adequate adjustment in any social, cultural, ethnic, or national surroundings. In most instances, however, the implications of psychological characteristics are likely to depend on their sociocultural context. For example, whether being a passive, dependent, group-oriented, and self-sacrificing person promotes effective functioning and a sense of well-being will depend on whether the person is part of a society that values these personality characteristics or that instead is embedded in a family, neighborhood, or societal context that values assertiveness, self-reliance, and ambitious striving for personal self-fulfillment.

3.4. Computer-Based Test Interpretation

Computer software programs are available for most standardized and widely used clinical assessment instruments. The data entry for these programs consists of the score for each test item or response, and the output typically includes (a) a list of these scores; (b) a set of indexes, summary scores, and profiles calculated on the basis of the test scores; and (c) a narrative report concerning the interpretive significance of these indexes, summary scores, and profiles. This narrative report is the key feature of what is commonly called computer-based test interpretation (CBTI).

Self-report inventories also lend themselves to a computerized test administration in which the test items and the alternative responses to them (e.g., "agree strongly," "agree," "disagree," "disagree strongly") appear on the screen. Computerized administration of

performance-based measures has not yet been shown to be practicable, however, owing to the fact that the response data are observational, infinite in their variety, and frequently interactive (in the sense that responses may require certain kinds of inquiry to clarify how they should be scored).

Computerized testing programs usually reduce scoring and calculation errors and also save time in recording and arraying the various indexes, summary scores, and profiles that provide the basis for test interpretation. Computerization also facilitates the interpretive process itself by virtue of a computer's complete scanning and perfect memory. A good CBTI program does not overlook any encoded implications the data are believed to have for psychological characteristics of the person being assessed.

On the other hand, CBTI is limited by its unavoidable nomothetic focus. Computer-generated interpretive statements describe features of a test protocol and present inferences that are applicable to people in general who show the same or similar test features. These statements do not necessarily describe the person being assessed, nor do they identify idiographic ways in which the person differs from most people who show certain similar test scores. Consequently, because of human individuality, CBTI narrative reports nearly always include some statements that do not apply to the respondent or that are inconsistent with other statements.

Accordingly, informed and professionally responsible computerized testing requires thorough knowledge of how and why test scores generate certain interpretive statements. Sophisticated clinical application of computer-generated test interpretation also requires appreciation of what a computer-generated report implies in the context of each respondent's personal history and current life circumstances. Unfortunately, CBTI programs can be obtained and used by persons with little or no knowledge of the tests on which they are based, especially in the case of self-report software that conducts the test administration, and with little appreciation of how and why to translate the nomothetic statements in computer-generated reports into the idiographic statements that are essential for an adequate clinical report.

4. FUTURE DIRECTIONS

Research findings have validated a large number and a wide variety of clinical assessment methods as providers of accurate and dependable information about the psychological characteristics of people. This information has, in turn, been shown to be helpful to mental health professionals in making decisions about differential diagnosis and treatment planning. However, little is yet known about the long-term impact of decisions made on the basis of clinical assessment findings. Are people better off as a result of having been assessed than they would have been otherwise? Do assessment-based decisions enhance the quality of life that people are able to enjoy? These questions set an important agenda for further directions in clinical assessment research that should emphasize follow-up studies of the long-term effectiveness of assessment procedures.

See Also the Following Articles

Attitude Measurement ■ Educational and Child Assessment ■ Forensic Mental Health Assessment ■ Intelligence Assessment ■ Measurement and Counseling ■ Mental Measurement and Culture ■ Personality Assessment ■ Psychological Assessment, Overview ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Psychophysiological Assessment

Further Reading

Beutler, L. E., & Groth-Marnat, G. (2003). *Integrative assessment of adult personality* (2nd ed.). New York: Guilford.

Butcher, J. N. (Ed.). (2002). *Clinical personality assessment* (2nd ed.). New York: Oxford University Press.

Butcher, J. N. (2003). Computerized psychological assessment. In J. R. Graham, & J. A. Naglieri (Eds.), *Assessment psychology* (pp. 141–164). Vol. 10 in I. B. Weiner (Ed. in Chief), *Handbook of psychology*. Hoboken, NJ: John Wiley.

Dana, R. H. (Ed.). (2000). *Handbook of cross-cultural and multicultural personality assessment.* Mahwah, NJ: Lawrence Erlbaum.

Maruish, M. E. (Ed.). (1999). *The use of psychological testing for treatment planning and outcome assessment* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Maruish, M. E. (2003). Psychological assessment in treatment. In J. R. Graham, & J. A. Naglieri (Eds.), *Assessment psychology* (pp. 119–140). Vol. 10 in I. B. Weiner (Ed. in Chief), *Handbook of psychology*. Hoboken, NJ: John Wiley.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland., K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128–165.

Rogers, R. (Ed.). (1997). *Clinical assessment of malingering and deception* (2nd ed.). New York: Guilford.

Weiner, I. B. (2003). The assessment process. In J. R. Graham, & J. A. Naglieri (Eds.), *Assessment psychology* (pp. 3–26). Vol. 10 in I. B. Weiner (Ed. in Chief), *Handbook of psychology*. Hoboken, NJ: John Wiley.

Cognition and Culture

Deanna Messervey and Li-Jun Ji

Queen's University, Kingston, Ontario, Canada

Yukiko Uchida

University of Michigan, Ann Arbor, Michigan, USA

GLOSSARY

analytic thinking The type of thinking that focuses on the object (instead of context) and the categories to which the object belongs. Analytic thinkers rely on rules and formal logic in understanding the world. This type of thinking is more prevalent among Westerners, particularly European North Americans.

correspondence bias (fundamental attribution error) The tendency to overemphasize a person's dispositions and to downplay the importance of the situation when explaining behaviors.

entity theory of intelligence The belief that intelligence is fixed and unchangeable.

holistic thinking The type of thinking that emphasizes context and relationships, and makes relatively little use of categories and formal logic. This way of thinking is prevalent among East Asians.

incremental theory of intelligence The belief that intelligence is mutable and can be changed.

independent self A type of self-construal, or self-concept, in which the self is perceived as a distinct entity that has a unique combination of traits, abilities, and other inner attributes. The independent self acts in accordance with his or her private thoughts and emotions rather than in a manner that is motivated by the thoughts and emotions of others.

interdependent self A type of self-construal in which the self is seen in relation to others, and the self is motivated to create and maintain interpersonal relationships with these relevant others. People with interdependent selves generally place the needs and goals of their in-group above their own.

perspective taking The ability to take the point of view of another person.

Culture refers to shared values, beliefs, and norms among a group of people who most often speak the same language and live in proximity. Cognition is defined as the process and products of human perception, learning, memory, and reasoning. Culture affects many aspects of cognition, and cognitive tendencies also guide the propagation of culture. This article provides an introduction to understanding how culture interacts with cognitive processes at basic and social levels from the perspective of cultural psychology. We discuss how culture affects the way that people perceive themselves, categorize information, and make social inferences. Moreover, we address how decision

making, memories, and communication and reasoning styles may be influenced by our cultural backgrounds.

1. SELF

In 1991, Markus and Kitayama outlined two types of self-concepts or self-construals: independent and interdependent. An independent self is perceived as a distinct entity who has a unique combination of traits, abilities, and other inner attributes. Moreover, the independent self acts in accordance with his or her private thoughts and emotions rather than in a manner that is motivated by the thoughts and emotions of others. It is worth noting that the study of the self has largely taken a "monocultural approach" that has originated in the West. Because North Americans predominantly have independent self-construals, it is not surprising that North American psychologists have adopted an approach that treats the self as an independent entity. When researchers, however, examined how the self is construed in other cultures, they discovered that independent self-construals were not as prevalent as in North America. Indeed, interdependent selves are found to be quite common among some other cultural groups, such as Asians. People with interdependent selves see themselves in relation to others and are motivated to create and maintain harmonious interpersonal relationships with these relevant others. Although people with interdependent selves hold internal attributes, they also understand the importance of context in shaping these inner attributes. Consequently, they are less likely to act in accordance with their thoughts, emotions, and other attributes, especially when such behavior affects relevant others. Thus, interdependent selves are guided by the way the self is related to others. According to Markus and Kitayama, interdependent selves are predominant in Chinese, Japanese, Hispanic, Filipino, African, and Hindu populations, and independent selves are predominant in Western populations.

The distinction between independent and interdependent selves followed Triandis's work on individualism and collectivism, which has received tremendous attention in the cross-cultural psychology literature. Individualist selves place their own personal goals above their groups' goals, whereas collectivists give priority to their groups' goals over their own. It is worth noting that research suggests that individualism and collectivism are not opposite dimensions. In 2002, Oyserman and colleagues conducted a meta-analysis examining individualism and collectivism across numerous cultures and found that people who are high in one dimension (e.g., collectivism) are not necessarily low in the other dimension (e.g., individualism). For example, European North Americans tended to be high in individualism yet were not less collectivistic than Japanese or Korean samples.

2. ATTENTION AND PERCEPTION

Culture affects how we perceive the world and where we focus our attention. Particularly, it shapes the extent to which we attend to the central object or to the field. In addition, differential attention may affect people's subsequent cognitive performances. To test cross-cultural differences in perception, researchers have used the rod-and-frame test (RFT) among other tests. RFT is made of a box (also called a frame) and a black line (also called a rod) at the end of the box; the frame and the rod can be turned independently of each other. When European American and Asian participants were asked to make perceptual judgments about the position of the rod, Li-Jun Ji and colleagues found that the position of the frame affected Asians' perceptions of the rod more than Americans' perceptions. In other words, European Americans were better than Asians at disassociating the object from the background and hence were more field independent. Conversely, Asians tended to make perceptual judgments based on the relationship between the object and the background and hence were more field dependent. In 1976, Berry argued that cognitive styles (such as field dependence and independence) are related to one's lifestyle. In particular, he found that sedentary agriculturalists were inclined to have a field-dependent cognitive style, and nomadic hunters and gatherers were more likely to have a field-independent cognitive style.

Similarly, Masuda and Nisbett presented European American and Japanese participants with an ocean scene on a computer screen. They found that Japanese participants paid attention to the background objects more than European Americans did (although Japanese participants did not necessarily pay less attention to the central object than did European Americans). As a result, Japanese participants were able to recall more details about the background than were European Americans. Such differential sensitivity to the background or the context has been found to facilitate performance on certain tasks but debilitate performance on other tasks. When the context remains constant, it

provides a helpful cue for Japanese participants to recognize the central object. When the context changes, however, it becomes difficult for Japanese participants to correctly identify the central object because of a possible cognitive binding between the central object and the background. On the other hand, changes in context do not have as much impact on European American participants, who generally pay little attention to the context anyway. Similarly, when a cognitive task requires disassociation between the object and the background, the task is easier for European North Americans than for East Asians. European North Americans are more likely to focus their attention on the central object and are better at ignoring the context in comparison to many Asians.

3. REASONING STYLES

For the past decade, Nisbett and colleagues have conducted a series of empirical studies on reasoning across cultures, which demonstrate that culture shapes people's thinking and reasoning styles in general. Specifically, many Westerners, particularly European North Americans, adopt an analytic system of thought. That is, they focus on the object (instead of context) and the categories to which the object belongs, and they rely on rules and formal logic in understanding the world. In contrast, East Asians tend to think holistically such that they emphasize context and relationships between objects. Moreover, they rely more on intuition and experience and make relatively little use of categories and formal logic compared to Westerners.

In a study conducted by Norenzayan and colleagues, Korean and European American participants were asked to rate the extent to which various arguments were convincing. Participants were presented with either typical arguments, (e.g., "All birds have ulnar arteries. Therefore, all eagles have ulnar arteries.") or atypical arguments (e.g., "All birds have ulnar arteries. Therefore, all penguins have ulnar arteries."). Koreans were more inclined to rely on their experiences and to indicate that typical arguments were more convincing than atypical arguments. On the other hand, European Americans, who tend to rely on formal logic and think analytically, saw both arguments as equally convincing. Similarly, when Korean and European American participants were asked to rate the logical validity of valid and invalid arguments that lead to either plausible or implausible conclusions, cultural differences emerged. In particular, Koreans were more likely

than European Americans to report that arguments that had plausible conclusions were more logically valid than arguments that had implausible conclusions.

According to Peng and Nisbett, differences in the reasoning styles between the East and the West also affect the way people think about contradictions. Chinese are much more likely than Westerners to seek compromising solutions when presented with opposing views. For example, European Americans who are told that (i) felines eat fish or (ii) felines do not eat fish often favor one argument at the expense of the other. That is, Westerners often report that one of the options is correct and the other is incorrect. Chinese respondents, however, tend to seek a solution that reconciles the two options. In this manner, Chinese tolerate apparent contradiction in an effort to find the "Middle Way." Similar cultural differences emerge when people attempt to resolve social conflicts. American and Chinese participants were presented with two scenarios that describe everyday conflicts (i.e., a conflict between a mother and daughter and a conflict that pitted focusing on school versus having fun). Peng and Nisbett found that, in general, Americans tended to favor one side of the conflict at the expensive of the other (e.g., mothers need to respect their daughters' autonomy more), and Chinese generally favored a compromising solution (e.g., both parties need to better understand each other).

4. CATEGORIZATION

In order to make sense of the world that is full of infinite information, we classify and categorize information in a way that can be culturally bound. For instance, Lopez, Atran, Coley, Medin, and Smith found that European Americans and Itzaj categorize mammals in ways that are similar in some situations yet different in others. The Itzaj are Mayans who live in the Peten rain forest region in Guatemala. They spend considerable time interacting with their environment, such as engaging in hunting activities or agriculture. In some regards, both European Americans and Itzaj have developed similar classification systems for mammals. For example, both cultural groups classify opossums in the same category as porcupines, even though opossums are marsupials and are in a class that is completely distinct from porcupines. In other respects, Itzaj are more likely than European Americans to classify mammals on the basis of ecological properties, such as behavior and habitat. For example, Itzaj classify

animals that dwell in trees together, whereas European Americans do not.

Deeper cultural differences emerge when people are asked to classify a variety of information. For example, European American participants tend to category things based on their category memberships, such as grouping "banana" and "orange" together because they are both fruits. Chinese children and Chinese University students tend to group things based on their functional relationships with each other, such as grouping "banana" and "monkey" together because monkeys eat bananas. According to Mishra, this latter type of grouping is also present in other cultures, such as the Birhor, Asur, and Oraon of India. Among European North Americans, however, this latter form of categorization is only found among children and is considered less sophisticated.

When categorizing information, East Asians rely more on intuitive experience, whereas European Americans rely more on formal rules. Generally, European Americans tend to assign objects to categories based on properties that fit the rule. For them, understanding, describing, and explaining objects involve classifying and seeking fundamental physical causes. Conversely, East Asians see an object as it relates to its context and the relationships among various objects. Because Asians historically did not generally perceive and classify objects as being distinct entities, they sought practical knowledge rather than developed formal rules or models.

5. INTERPERSONAL PERCEPTION AND MEMORY

Given that other people are a significant part of Asians' self-definitions, Asians also tend to pay more attention to other people's behaviors and emotions. They are expected to remember other people's preferences, read other people's minds, and anticipate the needs of others. This finding is related to the fact that Asians tend to be less direct than North Americans in expressing themselves.

Accordingly, European North Americans and Asians also differ in the way they remember information. Indeed, East Asians pay more attention to and remember more behaviors of their friends than do North Americans. When asked to recall and describe situations where they provided support to a friend, Japanese spend more time discussing the situations of their friends than their own actions. Conversely, European Americans describe their own actions more than their friends' situations. Furthermore, Japanese accounts of helping behavior tend to correspond more closely to their friends account compared to European Americans. Interestingly, there was little or no consensus between American participants and their friends regarding the nature of support that is exchanged in their relationship.

A special case of interpersonal perception is perspective taking. Perspective taking refers to the ability to take the point of view of another person. In a conversation, East Asians (such as Japanese and Chinese) sometimes identify themselves from another person's perspective. For example, a Japanese mother might call herself "Oka-san" (mother) in conversation with her child. The self is experienced through the other's perspective and his or her role in that relationship (e.g., mother–child). In other words, Asians tend to experience their own actions from the points of view of others. Similar differences have also been found in memory research. Research by Qi Wang has found that East Asians' memories tend to focus on other people or groups, whereas European North Americans' autobiographical memories tend to be self-focused and vivid. In one study, Asian North Americans were more likely to remember past episodes from a third-person's point of view, in which their memories included themselves. In contrast, European North Americans remembered the past episodes as they originally saw them; thus, the episodes they remembered tended not to include themselves. In addition to focusing on differential information when reporting autobiographical memories, European Americans tend to recount more emotional memories than do Chinese.

6. SOCIAL INFERENCE

Cultural differences are also found in social inference. Research suggests that Asians emphasize the context more than Westerners when explaining social events. For example, Asians tend to attribute the cause of behaviors to situational factors, whereas North Americans tend to attribute the same behaviors to personal factors. A well-known finding in North American culture is that people often fail to take contextual information into account when they judge the causes of others' behaviors. This phenomenon is called the correspondence bias, which is also known as the fundamental attribution error. In the classic paradigm,

participants are shown one of two types of essays: One essay expresses a favorable attitude toward a particular issue (e.g., capital punishment), and the other essay expresses an unfavorable attitude toward the issue. Participants are informed that the person who wrote the essay was able to choose which side of the issue he or she would like to discuss (i.e., choice condition) or was instructed to endorse a particular point of view (i.e., no choice condition). Next, the experimenter asks the participants to indicate the true attitude of the person who wrote the essay. Surprisingly, European Americans tend to judge the target person's attitude as corresponding to the view expressed in the essay, even when they know the target person did not have a choice. Consistent with an analytic thinking style, European Americans pay attention to the focal object (i.e., the essay writer) and erroneously believe that the person is independent from the situation. In contrast, Asians are more likely to recognize the situational constraints on the person's behavior when such constraints are made salient.

7. DECISION MAKING

Historically, there has been a dearth of research on how culture influences decision making. Recently, however, researchers have begun to give this area more attention. Much of this burgeoning research has focused on differences between East Asians and European Americans.

East Asians and European Americans tend to make decisions using different information. In one study, Incheol Choi and colleagues provided Korean and European American respondents with vignettes describing a target's prosocial and deviant behaviors. Participants were also given a list indicating possible reasons that might or might not explain the causes of the target's behaviors. Koreans considered generally more information than Westerners when asked to explain the target's behaviors. Likewise, other researchers have found that East Asians are more inclined to use historical information when making judgments about the present compared to Westerners. These cultural differences have been attributed to the different underlying theories people hold, in which East Asians have more complex and holistic theories than Westerners. In this manner, East Asians believe that many factors can interact to cause a particular outcome, whereas Westerners tend to believe that fewer factors can lead to a given outcome.

Culture can also affect the extent to which decision makers are willing to take risks. For example, Chinese are more likely to take financial risks than European Americans, although Chinese are more risk averse when they make medical or academic decisions. One possible explanation for these domain-specific patterns of risk taking is that Chinese have greater social ties (including family and friends) than European Americans, which can provide a "cushion" for any negative consequences associated with risk taking in financial domains.

Culture can also influence how confident an individual feels about his or her decisions or judgments. Typically, North Americans tend to make overly confident judgments. For example, after a general knowledge test, Americans are likely to judge that they are correct more often than they actually are correct. Yates and colleagues also found such overconfidence among Chinese, even though they did not find that Japanese made overly confident judgments. There are also different values associated with decision styles in different cultures. For example, decisiveness is not valued in Japan as much as it is in the United States or in China. These findings suggest that there are differences even between East Asian cultures, which call for further research.

8. LANGUAGE AND COMMUNICATION STYLE

Language is deeply interconnected with culture. On the one hand, language carries cultural meanings and values. On the other hand, language also transforms culture. Language functions as a tool that enables an individual to focus on certain aspects of the world. Research has found that Japanese pay greater attention to contextual factors during communication compared to European North Americans. For example, Japanese and Chinese do not use the first-person pronoun in some situations. When a teacher is talking to her students, she calls herself "teacher," and when she is talking to her child she refers to herself as "mother." In many English-speaking Western countries, however, the language focuses on the self, which promotes the idea that people are separate entities that are independent from their context. For example, in English people refer to themselves as "I" in almost every context. Moreover, English speakers usually do not drop personal pronouns when they are speaking.

Researchers have examined how language activates cultural identity. For example, Mike Ross and colleagues found that Chinese students in Canada reported interdependent self-perceptions when responding in the Chinese language. Interestingly, when asked to report their self-perceptions in English, they responded in a more westernized way. That is, they tended to see themselves as being more independent than interdependent. Likewise, Li-Jun Ji and colleagues found that Chinese bilinguals who learned English at a relatively late age categorized objects in a more relational (and Chinese) way when tested in Chinese than when tested in English. Thus, it appears that bilinguals' self-perceptions and cognition are closely related to language.

What is the role of language cross-culturally? According to Kim, talking is connected to thinking more in Western culture than in East Asian culture. When Asian Americans and Westerners were asked to think-aloud while solving reasoning problems, talking impaired Asian Americans' performances but enhanced the performances of European Americans. In the West, articulating one's thoughts and conveying such information to others enables the speaker to organize his or her thoughts. In contrast, verbal self-presentation is not encouraged among many Asians; thus, articulation of one's ideas does not occur naturally and actually hinders the reasoning process.

9. INTELLIGENCE

Jean Piaget had a considerable impact on the field of cognitive psychology and on intelligence research. He was the first psychologist to empirically measure the cognitive development of children over time. In this manner, he examined how the experiences of children are qualitatively different than adults. He advocated that the cognitive development takes place in four main stages that are universal regardless of culture (i.e., sensorimotor intelligence, preoperational, concrete operations, and formal operations). For instance, in the sensorimotor intelligence stage, children acquire mental representations, learn to imitate, and discover that objects continue to exist even when they are not in perceptual view. Much of Piaget's research has been criticized for a number of reasons. First, many sociocultural researchers do not support Piaget's universalist approach to cognition and maintain that development can take many routes rather than a single direction. Second, researchers contend that Piaget's notion of stages overlooks contextual factors that

shape an individual's development. Third, some researchers have argued that Piaget's stages are ethnocentric, in which people from cultures that do not develop in accordance with Piaget's stages are perceived as lacking in intelligence. Thus, these researchers see the theory as lacking generalizability. Although Piaget's research has played an important role in intelligence and cognitive development research, it appears that his theory is limited.

The history of defining and measuring intelligence has largely taken place in the West. When researchers began examining whether intelligence as defined by Western science could be applied to other cultures, they realized that other cultural groups often had different notions of what constituted intelligence. For example, the Trobriand Islanders may regard superior navigational skills as a key characteristic of an intelligent man. The !Kung of the Kalahari desert may define intelligence as the ability to spot sources of water. In China, intelligence may be viewed as knowing when and how to use language properly, such as knowing when (and when not) to speak. For Hindus, transcendental knowledge may be viewed as highly as practical knowledge. It appears that cultures define intelligence differently and that these differences stem from contrasting ecological and social experiences.

Not only can culture lead to diverse definitions of intelligence but also culture can influence the way individuals develop and maintain implicit theories concerning the malleability of intelligence. Dweck, Hong, Chiu, and colleagues argue that people hold either an "incremental theory" of intelligence or an "entity theory" of intelligence. Entity theorists believe that intelligence is fixed and unchangeable. For instance, European Americans tend to hold an entity theory. Consequently, they do not generally try to improve their performance on intelligence tasks when they have performed poorly because they generally believe that increased effort will have little effect on improving future performance. On the other hand, incremental theorists believe that intelligence is mutable and can be changed. East Asians tend to hold an incremental theory toward intelligence. Accordingly, they try to improve their performance on intelligence tests because they believe that increased effort leads to better performance.

Different definitions of intelligence make it difficult to develop a test of intelligence that can be used in all cultures. In addition, experience with Western cultures can sometimes have a considerable impact on test performance because most available intelligence tests

are developed in the West. For example, Australian Aboriginals who have more interactions with people from European backgrounds tend to perform better on standardized intelligence tests than Aboriginals from more isolated communities. In addition, some research suggests that performance on cognitive tasks is related to familiarity with testing materials. Children in Zambia often build intricate cars from wire, whereas children from England often draw pictures using pencils and crayons. Thus, Zambian children are more familiar with the medium of wire and English children are more familiar with paper-and-pencil mediums. Serpell found that Zambian children performed better than English children on a task that involved assembling a pattern using wire. Conversely, English children performed better than Zambian children on tasks that involved paper and pencils.

10. CONCLUSIONS

Research on culture and cognition has enriched our understanding of human behavior and has improved the quality of psychological theories by including people from non-Western cultures. This flourishing area of research has many great implications. Culture fosters people to think in different ways, and these differences may lead to misunderstandings and hostilities, from the long-lasting clashes between Jews and Arabs to the recent conflict between the U.S.-led coalition and Iraq. A world that appreciates and understands people from diverse backgrounds may help resolve existing and future conflicts.

Knowledge of culture and cognition may also help to smooth interactions at interpersonal levels. With increased globalization, the world becomes smaller and cross-cultural interactions become more frequent. Understanding how your colleagues from other cultures may view the world and reason differently than you do will help you to enhance your multicultural partnerships and collaborations. Examples of failing to recognize cultural differences in cognition can be found in business. For example, in a joint-venture company, making internal attributions for success by a

European American could be perceived as arrogant by a Chinese, whereas making external attributions for failure by a Chinese might be perceived as irresponsible by the European American. Ignorance of each other's thinking styles is likely to lead to negative interpersonal experiences and could even have negative consequences on the business outcome.

Understanding cultural differences in cognition can also play an important role in education. For example, teachers who deal with students from different cultural backgrounds would benefit from knowing more about how culture influences cognition. Some teachers may encourage students to verbalize their thoughts as a means of helping them develop their ideas. Although this approach to teaching may benefit Western students, students from other cultural backgrounds may actually be hindered by such an approach.

Understanding the relationship between culture and cognition enhances our understanding of the world, and we can benefit from such knowledge in different aspects of our lives.

See Also the Following Articles

Attention ■ Cross-Cultural Psychology, Overview ■ Decision Making ■ Intelligence and Culture ■ Interpersonal Perception

Further Reading

Berry, J. W., Dasen, P. R., & Saraswathi, T. S. (Eds.). (1997). *Handbook of cross-cultural psychology* (2nd ed., Vol. 2). Boston: Allyn & Bacon.

Berry, J. W., Irvine, S. H., & Hunt, E. B. (Eds.). (1988). *Indigenous cognition: Functioning in cultural context*. Dordrecht, The Netherlands: Martinus Nijhoff.

Lonner, W. J., & Malpass, R. S. (Eds.). (1994). *Psychology and culture*. Boston: Allyn & Bacon.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently ... and why*. New York: Free Press.

Triandis, H. C. (1995). *Individualism & collectivism*. Boulder, CO: Westview.

Cognitive Aging

Douglas H. Powell

Harvard Medical School, Boston, Massachusetts, USA

GLOSSARY

cognition (or "intellectual functioning") Global intelligence or a mental capacity that consists of several specific abilities; an IQ (for intelligence quotient) test such as versions of the Wechsler Adult Intelligence Scale (WAIS) are examples.

cross-sectional studies A type of research that studies age-associated changes in different individuals who are grouped by age; for example, the decline in cognitive functions might be measured over the life span by comparing the scores of individuals in age groups of 20, 40, 60, and 80 years.

crystallized abilities (or "verbal" or "left hemisphere" aptitudes) Aptitudes that are largely dependent on acquired knowledge and formal education and that develop rapidly during the first 20 years of life; vocabulary is an example.

dedifferentiation The relative decline in specific aptitudes from decade to decade that is nested within a greater decline in global intellectual functioning in normal older adults.

fluid abilities (or "performance" or "right hemisphere" aptitudes) Aptitudes that are largely independent of acquired knowledge and education and that are influenced by biological factors and the integrity of the central nervous system; block design is an example.

immune system The body's complex and multidimensional defense against invading bacteria, viruses, other pathogenic microorganisms, and tumors; it also assists in the repair of damaged tissue after injury and creates specific antibodies to combat diseases.

longitudinal studies A type of research that studies age-related changes in the same participants who are reexamined several times over their lifetimes.

MicroCog A computerized test of intellectual functioning that consists of 13 subtests measuring both crystallized and fluid abilities; it was administered to 1002 physicians (ages 25–92 years) and 561 normal respondents.

neuroimaging Imaging technology that enables investigators to observe how the brain functions when it carries out specific intellectual tasks; single photon emission computed tomography (SPECT), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) are examples.

optimal cognitive agers (OCAs) Those older individuals whose cognitive test scores are statistically similar to those in midlife.

psychoneuroimmunology The study of how the central nervous system is influenced by psychological and social factors.

Public Law 99-592 The law that eliminated age-based mandatory retirement for workers in the United States; it became effective January 1, 1994, and excludes only a small number of occupations (e.g., Federal Bureau of Investigation agents, airline pilots).

variability (or "variance" or "standard deviation") How widely dispersed, from top to bottom, the individual scores within an age group may be.

Interest in cognitive aging has grown substantially during the past 25 years. For instance, the divisions of the American Psychological Association (APA) expanding most rapidly during this period have been those with a strong interest in aging such as aging and neuropsychology, a field that examines cognitive impairments associated with aging (e.g., Alzheimer's disease). Responding to the growing interest in Alzheimer's disease and other conditions afflicting numbers of older people, the U.S. Congress designated the 1990s as the "Decade of the Brain." It might be said that that this interest applies especially to "older" brains. Funds are being invested at record high levels to increase the understanding and amelioration of the effects of age on intellectual vigor as well as on physical health. These efforts may be more important today than ever before. The reasons for this are demographic, legislative, physical, and economic.

1. INTRODUCTION

1.1. Demographic Trends

America is graying. At the start of the new millennium, one of every seven U.S. citizens was over 65 years of age. By 2025, this proportion will be one of every five. The group of those age 85 years or over, currently numbering in excess of 3 million, is the fastest-growing segment of the population. This increase in the numbers of elders is not a uniquely American phenomenon. By 2025, those age 65 years or over in Japan will double. The People's Republic of China, Korea, and Malaysia anticipate a tripling of this age group. There is little benefit in a longer lifetime if the quality of health and mental acuity are greatly diminished during these added years. The budget of the National Institute on Aging has tripled during the past decade. Substantial portions of that agency's funds are earmarked for research to reduce the incidence of Alzheimer's disease and other forms of cognitive impairment and to enhance intellectual vitality during the later years.

1.2. Legislative Changes

Compelling workers to retire due to age was banned in the United States by Public Law 99-592 as of January 1, 1994. With the exception of a small number of occupations, such as Federal Bureau of Investigation agents and airline pilots, this law eliminated age-based mandatory retirement for nearly everyone. A primary reason for passage of this law was that many studies had found that a large number of older individuals can match the physical and mental vigor of their younger colleagues. Because variability increases with age, it is necessary to differentiate between functional age and chronological age among seniors. Currently, organizations and professions are struggling with the implementation of objective and fair methods for retaining their older physically and cognitively vigorous employees while creating scenarios for encouraging the retirement of those whose abilities have declined to the point that their job performance is compromised.

1.3. Increased Physical and Cognitive Vigor

Despite the growing trend toward obesity in the United States, the physical and cognitive vigor of a large proportion of 60-, 70-, and even 80-year-olds today is superior to that of their parents and grandparents of the same age generations ago. Research confirms these observations. A 1997 report by the National Academy of Sciences showed a dramatic increase in the proportion of Americans over 65 years of age who are able to care for themselves. Reasons offered to account for these gains include reductions in drinking and smoking, better diets, weight loss, control of blood pressure, and use of aspirin to reduce heart attacks and strokes. Also, modern medicine has been chipping away at physical conditions that greatly affect cognition, including medications for diabetes, hypertension, and heart problems as well as advances in cardiac surgery and neurosurgery. Better physical health is strongly correlated with higher levels of intellectual functioning among older individuals.

1.4. Economic Needs

One unanticipated benefit of Public Law 99-592 is that the American economy will need its aging "baby-boomers"—those born during the 1946–1964 era—to work well past what used to be the normal retirement age into the first quarter of the 21st century. The reason is that the labor pool is shrinking due to the "baby-bust" generation—those born during the 1965–1977 era—now entering the workforce. Because their yearly birth pools were much smaller than those during the previous 20 years of the baby boom, the labor force will experience a shortfall in those age groups that organizations have traditionally relied on to form the backbone of their workforces, that is, those 30 to 44 years of age.

By 2009, there will be 10.6% fewer men and women in that age group than there were a decade earlier. Census projections estimate that there will not be as many people in that age group as there were in 2000 until 2025. A source of workers that could fill this projected labor shortage are men and women in their 60s and early 70s. Fully 70% of the baby boomers, now at midlife, say that they want to stay on their jobs after reaching 65 years of age. A large proportion of these individuals are mentally and physically fit to continue working if opportunities are made available.

2. COGNITIVE CHANGES ACROSS THE LIFE SPAN

What happens to cognitive abilities over the life span? Which aptitudes are impaired most dramatically, and which are most likely to be spared? How much variability is there among people of a given age as they grow older? This section examines global changes in cognition through the life cycle and trace the age-associated pattern of decline among particular aptitudes, paying close attention to the degree to which some are impaired or spared. It also discusses evidence in support of increasing dedifferentiation of cognitive functions with advancing age.

2.1. Overall Cognitive Changes through the Life Span

Figure 1 shows the decline in Full Scale IQ from 30 to 75-plus years of age on an earlier version of the Wechsler Adult Intelligence Scale (WAIS). The curve is derived by comparing the average of the scaled scores for each older age group decade with the norms for 25- to 34-year-olds. Overall, cognitive scores slope gently downward until 60 years of age. At that point, the level of intellectual functioning of the average 60-year-old is only 10% less than that of the average 30-year-old. But then the cognitive decline accelerates. During the next decade, the overall IQ drops 9%. The scores of those age 75 years or over are 11% lower than those of 70-year-olds.

A limitation of these data is that the WAIS respondents were not matched for characteristics (other than age) that affect cognitive changes over the life span. These other characteristics are educational level, social class background, and access to health care. Those who have more education, are wealthier, and are able to avail themselves of adequate medical services usually score higher on cognitive tests than do their less fortunate age mates. Studies that have rigorously matched age groups on these characteristics have found that the shape of the curves of overall cognitive decline is similar but that the test scores are higher at every age for those with more advantages.

2.2. Decline of Specific Aptitudes through the Life Span

Although overall intellectual functioning declines predictably from 30 years of age onward, the downward progression of individual aptitudes varies greatly. For instance, the various versions of the WAIS all have found that the Verbal IQ scores slope downward far more gradually than do the Performance IQ scores. For
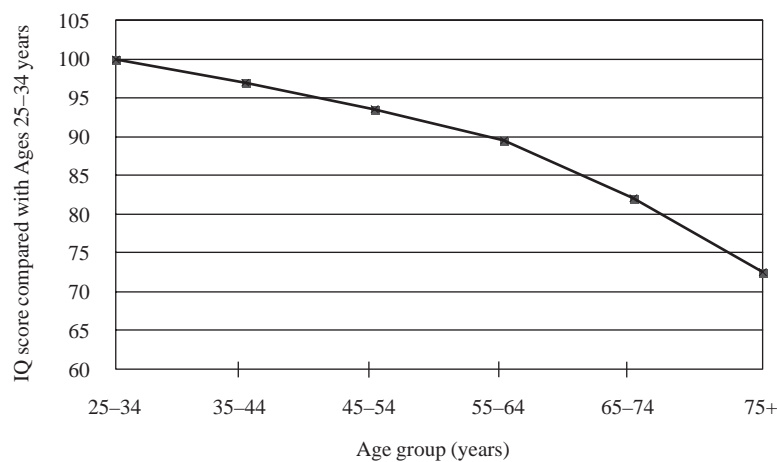

FIGURE 1 Average WAIS Full Scale IQ scores for older age groups compared with 25- to 34-year-olds. Adapted from Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence.* Boston: Allyn & Bacon.

instance, scores on the Verbal IQ decline 10% from 30 to 70 years of age, whereas scores on the Performance IQ drop 25% during this same period.

The types of abilities in the WAIS Verbal and Performance sections explain why there is so much difference in the rate of decline. The Verbal IQ score is based largely on crystallized abilities. In contrast, the Performance IQ subtests are composed of largely fluid aptitudes. Research by Salthouse and colleagues has demonstrated that fluid aptitudes, such as processing speed and working memory, are greatly impaired by aging. Tests involving processing speed typically require solving complex problems under time pressure. Being required to learn a number–symbol code and then to enter the proper number next to a list of symbols is an example. Working memory involves storing an important piece of information while working on another task. Backward number recall is an example.

Age differences in processing speed and working memory apply much more to complex tasks than to simpler ones. For instance, little difference occurs between younger and older adults on analogies of low and medium difficulty under time pressure. Where the performance of older test takers falls off, relative to that of young adults, is on the hardest of the analogies (e.g., "Fission is to splitting as fusion is to ____") under stressful time limits. The differences are shown in Fig. 2.

Research reports summarized by Heaton and colleagues in 1981 show other aptitudes whose scores decline earlier and more steeply with increasing age. Timed spatial recall, reasoning, and verbal memory after delay are examples. In contrast, attention, word knowledge, and calculation skills remain relatively stable and are largely spared by the aging process.

All age groups have more difficulty with tasks requiring dual-task or divided attention. The increased incidence of automobile accidents among drivers using cell phones is an example. As people age, their dual-task performance declines rapidly. Over the past three decades, scientific studies have consistently reported that younger adults outperform older adults on many different tests of dual-task attention, for example, classifying a list of words appearing on a computer screen as verbs or nouns while simultaneously listening to a string of numbers and pushing a button when hearing two odd numbers in a row.

Although there is clear evidence of substantial disparity in the rate of decline of specific aptitudes, a number of studies have reported data showing dedifferentiation of cognitive functions with advancing age. That is, in normal older adults, the decline in specific aptitudes from decade to decade is embedded within a greater decline in global intellectual functioning. The evidence is stronger in cross-sectional data and among test scores based on accuracy alone. Longitudinal studies and test results combining speed with accuracy show weaker but similar patterns. Explanations for this phenomenon vary. One theory is that dedifferentiation is simply a symptom of cognitive decline. Another view, based on imaging studies, is that older adults compensate for diminished skills by drawing resources from other areas of the brain that are not normally activated during mental tasks.

2.3. Alzheimer's Disease

If the 1990s was indeed the Decade of the Brain, it could also be said that the 1980s was the decade when Alzheimer's disease (AD) and other forms of dementia began to attract public attention as potential
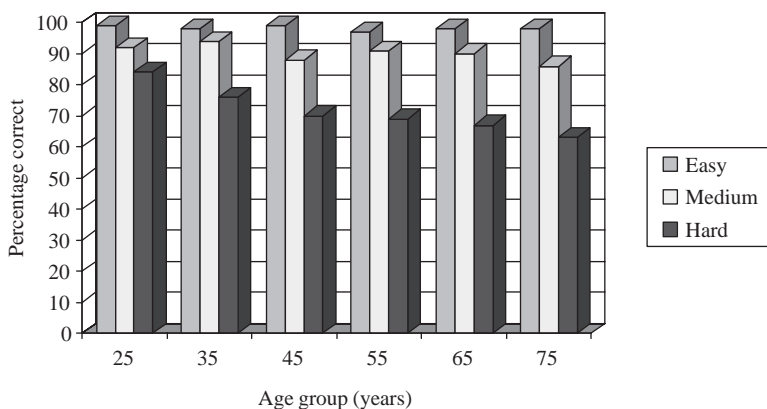

FIGURE 2 Effect of time pressure on analogy scores by difficulty level. Adapted from Salthouse, T. A. (1992). Why adult age differences increase with task complexity. *Developmental Psychology, 28,* 905–919.

threats to those entering the decades beyond middle age. The Alzheimer's Disease and Related Disorders Association was founded in 1980. In that same year, according to Lexis/Nexis (an electronic database for media publications), major U.S. newspapers mentioned AD or dementia twice. In the year 2000, there were 250 separate pieces about AD or dementia.

AD was first described by a German psychiatrist, Alois Alzheimer, in 1907. His first case was a 50-year-old woman whose memory and other cognitive functions deteriorated rapidly over a 5-year period before she died. An autopsy found abnormal brain structures that are referred to as neurofibrillary tangles. and neuritic plaques. As this case demonstrates, the relatively young onset differentiates AD from senile dementia. Today, health care specialists distinguish between presenile AD (prior to 65 years of age) and senile AD (age 65 years or over).

The criteria for AD include a gradual decline in at least two intellectual functions (e.g., attention, memory, language, reasoning, spatial ability) that are sufficient to significantly impair social relationships and/or performance at work. This decline cannot be explained by head trauma, neurological conditions, ill health, severe psychiatric symptoms, or drug/alcohol abuse. The cognitive deterioration also may be associated with depression, emotional outbursts, and/or apathy. It is not unusual to find AD occurring in individuals who otherwise are in good health.

At the start of the new millennium, there were an estimated 5 million people in the United States with AD or other forms of dementia. Prevalence estimates vary, however, and depend on the criteria used for making the diagnosis. For instance, a research team in England rated 100 elderly, community-dwelling volunteers using seven different sets of diagnostic criteria for AD. They found that between 3 and 63% of this population would be judged as "having" AD, depending on which standards were applied. A survey of the incidence of AD among older people in seven countries found proportions ranging from 1.9 to 52.7% of the population. One explanation for the disparate findings on the frequency of AD is the differences in the way in which the condition was assessed. Diagnoses that used tests that included measures of both crystallized and fluid abilities produced more consistent estimates. Applying these test-based standards, AD appears to afflict approximately 3% of 65- to 69-year-olds, 6% of 70- to 74-year-olds, and 11% of 75- to 79-year-olds. Beyond 80 years of age, the proportion of AD rises sharply.

3. AGE-RELATED VARIABILITY

Studies of age-related changes in cognitive functions have concentrated on what are called "measures of central tendency," which are most often the mean or average scores. Figures 1 and 2 are illustrations of this approach. Typically, the mean scores of 30-year-olds are compared with the averages of groups of those in older decades. The downward slope of these age group means is then reported. What has been overlooked by focusing on averages alone is what happens to the variability of the individual scores with each advancing decade. The increasing variability with age found in the cross-sectional study of the cognitive aging of 1002 physicians using a computerized test, MicroCog, is an example.

Figure 3 shows the percentage of decline in MicroCog total score in each age group after 40 years. This is compared with the percentage growth of the variability for each decade during this same period. It can be seen that the standard deviations of the age groups rise much more rapidly than the mean scores decline. At 70 years of age, the average physician's MicroCog total score was 13% lower than that of the typical 40-year-old. But the variability of the scores for these 70-year-olds was 57% greater than that for their younger colleagues.

Growing variability within advancing age groups is not specific to these data. Summaries of other studies of age-associated changes in cognition have reported that four of every five studies found that variance within each age group increased with advancing age. These findings occurred equally in both cross-sectional and longitudinal research.

The practical value of being aware of this growing age group variability is that it indicates that a number of older individuals continue to function as well as those in their prime. These older individuals might be called optimal cognitive agers (OCAs).

4. EVIDENCE FOR OPTIMAL COGNITIVE AGING

Newspapers and magazines often features stories of elders who have been productive and successful until very late in life. Novelists James Michener and Barbara Cartland were still writing into their 90s. Heart transplant pioneer Michael DeBakey, South African political leader Nelson Mandela, musician Lionel Hampton, and humanitarian Mother Teresa worked productively in their 80s. Grandma Moses did not begin painting until
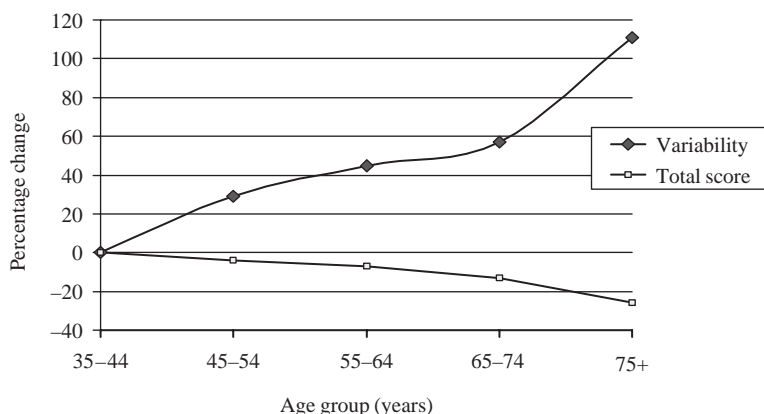
FIGURE 3 Percentage changes in MicroCog total score and variability from 40 years of age: 1002 Physicians. Reprinted by permission of the publishers from *Profiles in Cognitive Aging* by Douglas H. Powell in collaboration with Dean Whitla, Cambridge, MA: Harvard University Press, Copyright © 1994 by the President and Fellows of Harvard College.

she was 78 years old and continued well beyond her 100th birthday.

Social scientists are sometimes skeptical of such examples, wondering whether these people are merely statistical anomalies rather than representative of the vast majority of their age mates. However, empirical evidence now confirms the presence of large numbers of elders who continue to function cognitively at a level similar to those in midlife. This group of individuals, the OCAs, can be distinguished from their contemporaries by applying statistical standards.

Consider the following example. The MicroCog study described earlier obtained cognitive test scores on a large number of physicians from 25 to 92 years of age. Suppose that we wanted to know what proportion of older physicians were OCAs. We could use those doctors in their prime (say, 45–54 years of age) as a comparison group and then set a statistical threshold for optimal cognitive aging. For example, we might choose a cutting score of the 10th percentile, approximately −1.3 standard deviation, of those 45 to 54 years of age. In other words, for doctors to qualify as OCAs by this standard, their scores on MicroCog must be higher than the bottom 10% of those in midlife. The percentages of physicians age 65 years or over who meet this statistical standard are shown in Table I.

These data confirm what informal observation has told us: A large number of people continue to function at a high level cognitively well beyond 65 years of age. Nearly three-quarters of the 65- to 69-year-old physicians, and half of the 70- to 74-year-old physicians, scored as well on this cognitive test as did those in their prime. One-quarter (25%) of 75- to 79-year-old physicians, and

TABLE I
Proportions of Older Physicians with MicoCog Scores above the 10th Percentile of 45- to 54-Year-Old Physicians

Age group	Percentage above 10th percentile
65–69 years	73.2
70–74 years	50.0
75–79 years	25.0
80+ years	10.4

Source. Adapted from Powell, D. H., in collaboration with Whitla, D. K. (1994). *Profiles in cognitive aging.* Cambridge, MA: Harvard University Press.

slightly more than 10% of those beyond 80 years of age, performed as well as the reference group. These data are not specific to physicians. Similar proportions were found among nonphysician respondents in this study.

Although there are advantages to using respondents matched for education, social class, and access to health care for these calculations, nearly any cognitive test data set that reports individual test scores by age group through the life span can be analyzed to yield the percentage of older participants who are OCAs.

5. CORRELATES OF OPTIMAL COGNITIVE AGING

What distinguishes these OCAs from their contemporaries whose cognitive decline follows the average downward trend for their age group? Are there behaviors

within individuals' control that are correlated with optimal cognitive aging? If so, identifying these factors may contribute to behavioral changes that can improve the quality of these later years for many people.

But first, another question must be asked: What are the relative contributions of heredity and environment to the quality of cognitive aging? Research finds that heredity does play a strong role in the acquisition, maintenance, and decline of most human qualities. It is also true that environmental forces exert a strong effect. Swedish scientists tested 80-year-old identical twins who had been separated at birth and raised in different home environments. The researchers estimated that approximately 60% of the variance (i.e., factors contributing to the similarity of cognitive ability) was genetic and that the rest was related to environmental influences. Reports from dozens of other twin studies have estimated that heredity accounts for 50 to 60% of what goes into mental ability. So, it seems that somewhere between 40 and 50% of the factors influencing the quality of cognitive aging is related to environmental influences, that is, factors that are partly within the control of each individual.

This section summarizes those behaviors that are correlated with optimal cognitive aging and that are within individuals' control to varying degrees. These include activities that have been demonstrated to be directly related to higher levels of cognition in older adults. Also, this section describes a number of activities that indirectly benefit cognition because they are associated with better physical health. Finally, the section discusses selective optimization with compensation as a strategy for minimizing the inevitable negative effects of cognitive aging.

5.1. Activities That Directly Benefit Cognition

Activities that directly benefit cognition include regular exercise, mental enrichment, active social networks, and cognitive training. Of all the activities that could be recommended to maintain cognitive vitality, regular exercise has the greatest positive effect. This is because exercise strengthens the cardiovascular system, which maintains necessary blood flow to the brain. Cardiovascular disease is a major cause of cognitive impairment among individuals of any age. Meta-analyses of research over the past 40 years have found that younger and older individuals working out for 30 minutes or more every other day for at least 3 months perform at a higher

level on tests of attention, memory, spatial ability, processing speed, and reasoning than do their sedentary neighbors. Of interest is that those individuals whose regimen combines strength/flexibility and aerobic training obtain slightly higher cognitive scores than do those who engage in aerobic or strength/flexibility exercise alone.

Physical and mental vigor can be improved until quite late in life. An illustration of the power of these programs was strength training for 90-year-olds at the Hebrew Rehabilitation Center for the Aged in Boston. After just 8 weeks of regular workouts, these nonagenarians increased their muscle strength by 174% and increased their walking speed by 48%. Although pre- and postcognitive tests were not given, it is likely that improved intellectual vitality accompanied the physical gains. Baylor Medical School researchers found that physically active women and men in their late 60s had greater cerebral blood flow that resulted in higher cognitive test scores than those of their inactive contemporaries.

Mental enrichment stimulates cognitive functions. K. Warner Schaie, the director of the Seattle Longitudinal Study, and his coworkers identified factors that they believe are correlated with maintaining high levels of mental vigor well into later life. These include being open to new and stimulating experiences and having close relationships with individuals who have strong intellectual abilities. Scientific support for this thinking has come from the laboratory of Marian Diamond at the University of California, Berkeley. She was among the first to discover that laboratory animals that were raised in a more complex and challenging environment, along with other animals, developed a thicker frontal cortex than did those that lived in normal but unstimulating cages. She discovered that this occurred in both young and middle-aged animals. The increased thickness was because the existing brain cells in the cortex produced more connections (called dendrite branching) in response to outside stimulation.

A research team at the University of California, Los Angeles, confirmed these findings with humans. The researchers looked at the brains of 20 adults at autopsy. Then, they correlated the size of the nerve cells in the brains' cortexes with the activities reported by these individuals prior to death. They found that the cortexes of those older people who were more mentally and socially active were thicker than those of individuals who were not.

Social support has long been known to be correlated with physical health, but until recently little attention had been paid to its effect on cognition. Scientists from the MacArthur Studies of Successful

Aging studied the impact of social networks on a battery of aptitude tests given to 70- to 79-year-olds. They found that those who had more emotional support obtained higher test scores. At follow-up testing nearly 8 years later, higher retest aptitude scores were moderately correlated with having had greater social support. Other research has found that having larger and more diverse social networks and having more frequent contact with friends and family are related to less cognitive decline.

The question of whether age-related cognitive decline can be reversed through training has raised the interests of specialists in gerontology for the past quarter-century or so. Overall, it appears that most normally functioning individuals in the "young–old" years (approximately 60–75 years of age) and many of those who are older have the capacity to enhance their mental capabilities through practice. One example is the improvement made by participants in the Seattle Longitudinal Study who received 5 hours of training when they were 67, 74, and 81 years old. On average, the two younger groups improved their reasoning and spatial abilities to the levels of their test scores of 14 years earlier. Octogenarians were able to improve their performance to the level of 7 years earlier.

Follow-up studies of cognitive training have found that only those mental skills selected for remediation improved. Those not targeted for tutorial work remained the same as when the instruction began. Thus, if the goal of the program was to better remember a list of words, the participants did not wind up also exhibiting stronger calculation skills at the end of training.

Surveys of the published reports on cognitive training have found that many different methods produce positive results. Older participants at the Max Planck Institute in Berlin who spent 5 hours studying self-help books on improving fluid (spatial) abilities performed as well on tests as did those who had 5 hours of training. These gains were maintained on tests 6 months later. Similar results have been achieved with a wide range of other interventions, with some of them being quite indirect. For instance, members of groups that practiced transcendental meditation regularly improved their scores on cognitive tests far more than did those who did not.

So far, the effects of the training are largely untested in real life. After completing a course on memory improvement, is a woman better able to recall a list of errands she must run? After cognitive training, can a man more easily remember where he left his car in a supermarket parking lot?

5.2. Activities That Indirectly Benefit Cognition

Intellectual functions are indirectly influenced by a number of activities that enhance physical health. In an otherwise reasonably healthy aging population, these include behaviors that benefit cardiovascular functions and strengthen the immune system.

In addition to regular exercise, two crucial contributors to cardiovascular health are weight control and blood pressure management. Obesity is a growing problem in the United States. Depending on which criteria are applied, the proportion of overweight adults in the United States currently varies from 37 to 67%. By any standard, the percentage of obese adults has nearly doubled during the past 40 years. Excess weight is associated with greater risk for diabetes, hypertension, and cardiovascular disease—three conditions that adversely affect cognition. Obesity increases with age, partly because fewer calories are required to support normal activity with advancing years and partly because many older adults are physically inactive. In the year 2000, 35% of American older adults reported that they had engaged in no physical activity during the previous month.

Controlling blood pressure reduces the risk of cardiovascular disease and stroke. Participants in the Framingham Longitudinal Study with untreated hypertension scored lower on cognitive tests than did those with normal blood pressure. Increased diastolic blood pressure had particularly negative consequences for intelligence. The intellectual consequences of high blood pressure are far more serious for younger people. Hypertensives under 40 years of age had far lower scores on cognitive tests of attention, memory, mental flexibility, and reasoning than did age-matched normal individuals. Beginning to control high blood pressure early has substantial cognitive and physical advantages.

Physical vitality and cognitive vitality are greatly affected by the efficiency of the immune system. The immune system declines with age. Just as with age-related cognition, however, considerable variability occurs in immune system functions among older people. Until the past decade, it was assumed that the immune system functioned independently of psychosocial influences. Over the past 10 years, however, evidence has accumulated that the immune system is significantly influenced by mental states. This has led to increasing scientific interest in what is called psychoneuroimmunology. Of particular interest is the influence of stress on the suppression of the immune

system given that this, in turn, can adversely affect health and cognition.

How much stress individuals confront, how well this stress is managed, and what other resources individuals can draw on to ameliorate the impact of negative events and bring comfort into their lives all are found to be correlated with health status. For example, more than 400 adult volunteers in Pittsburgh, Pennsylvania, were given nose drops containing a common cold virus. Not surprisingly, a large number of the volunteers developed severe colds soon afterward. But not everyone developed colds, and the severity levels of the colds varied widely. These scientists also measured the total amount of stress that these individuals had to deal with during the past year. Then, they compared these data with the susceptibility to, and the severity of, the colds. They found that individuals with more stress in their lives developed colds far more often, and with more severity, than did those whose stress scores were below average.

Elevated stress-related emotions of anxiety, depression, and anger correlate with lower scores on mental tests. This may be because high levels of stress diminish intellectual resources. An illustration is the finding that neuronal loss in the hippocampus area of the brain was moderately correlated with the number of months in combat in the war in Vietnam. The negative effect of stress emotions appears to be greatest on fluid abilities. For example, higher levels of anxiety have been demonstrated to be associated with lower scores on aptitudes such as analogies, picture memory, and block design.

5.3. Selective Optimization with Compensation

One of the most promising strategies for minimizing the impact of aging on cognitive performance is selective optimization with compensation. Developed by Baltes and coworkers, the essential ingredients in this approach are broken down into three steps. The first step is selection, that is, giving up some activities to concentrate the available mental energies on those remaining activities. The activities that continue to have priority are usually selected due to skills, motivations, and opportunities. An aging physician who wishes to remain active might be an example. In midlife, the physician has a busy practice and also teaches medical students. He simultaneously does clinical research, serves on professional committees and boards, and does consulting work. By his early 60s, the physician may begin to focus his mental energies on those patients and types of activities that he likes best, cutting back on those professional responsibilities in which his interest is waning. So, the physician keeps working at his practice but takes no new patients, and he continues his teaching and research but gives up committee work and consulting. At 70 years of age, the physician may limit his professional activities to teaching and research.

The next ingredient is optimization. This means doing everything possible to maximize one's performance at those things that one chooses to do by anticipating the potential negative effects of cognitive aging. For example, the physician may begin to limit his practice to those patients who he believes he can treat most effectively, referring the others to colleagues. Keeping up with the recent developments in medicine and taking short courses to maintain and upgrade his medical skills become priority items. The physician may try to work "smarter," that is, using computerized expert analyses and digital libraries more readily as opposed to trying to keep all of this information in his head. Optimization requires more time for preparation. For an older physician, this might entail carefully reviewing the chart of a patient in advance of an appointment rather than counting on his ability to get a picture of the medical history by a quick scan while seeing the patient. The aging physician also begins to dictate notes right after seeing the patient instead of waiting until the end of the day or later in the week to do so. In his 60s, the physician must accept the reality that far more time needs to be set aside to prepare lectures than was the case a decade earlier.

Compensation is doing everything that can be done to bolster fading abilities. Eyeglasses and hearing aids, to see and listen as well as possible, are examples from the physical standpoint. From the cognitive standpoint, compensation can include doing one task at a time and recognizing the effect of circadian rhythm (day–night cycle) and glucose/caffeine on cognition. The still working older physician compensates by avoiding dual-task activities such as making notes in a patient's chart while talking about another patient on the telephone. Cognitive compensation might include putting to use an understanding of his personal circadian cycle. The average older person's mental functions are stronger in the morning than later in the day, so the physician may schedule his most difficult patients in the morning and schedule less challenging professional activities in the afternoon. Finally, the physician may compensate for diminished mental acuity in the late afternoon with a cup of coffee and a cookie. For most healthy older individuals, moderate amounts of caffeine and glucose enhance attention and memory for a short time.

6. LOOKING AHEAD

Judging from the remarkable progress and growth of interest in all aspects of aging during the past quarter-century, there is little doubt that this will continue to be an area of expanding scientific and academic interest, as well as an area of considerable opportunity, during the decades ahead.

Several relatively recent developments make cognitive aging an especially interesting subspecialty within the field of aging. For instance, developments in neuroimaging technology over the past decade or so have enabled investigators to see in real time how the brain functions when carrying out specific intellectual tasks, to visualize the differences between the brain structures of younger individuals and those of older individuals, and to observe how the brain is compromised by conditions such as AD. Neuroimaging techniques have made possible a quantum leap in understanding the relationship between the brain and behavior. A new discipline, called cognitive neuroscience, draws deeply from this technology. At the time of this writing, the field is attracting the interest of the some of the most influential thinkers in the social sciences and medicine.

Another important development has been the availability of findings from longitudinal studies of human development through much of the life span. Examples of these longitudinal studies include the Oakland/Berkeley Growth and Guidance Studies, the Terman Study of Children With High Ability, the Grant Study of Adult Development, the Seattle Longitudinal Study, and the Duke Longitudinal Studies. Some of these were begun during the 1920s and 1930s, whereas others started during the third quarter of the 20th century. Although many of these studies have limited generalizability because their participant pools are not representative of the U.S. population in terms of gender, education, social class, and race, they still provide an increased understanding of the aging human body and mind. As the participants from these investigations are moving into their later years, this research has begun to bear fruit. The numerous books and articles published about these studies have contributed much to the understanding of cognitive aging.

But these publications do not nearly exhaust the opportunities to learn much more about cognitive aging. The raw data from many of these longitudinal studies—the actual responses from each participant every time he or she was examined—are currently archived and are available for analysis. For instance, the Henry A. Murray Research Center for the Study of Lives at Harvard University contains nearly 300 complete data sets from these longitudinal studies as well as other cross-sectional research that is better balanced for demographic variables. The rapid growth in data storage technology and computer memory and speed makes it possible today to access the raw data from these aging studies and to examine them using a computer in the privacy of a distant laboratory, professorial office, or student's room. Today's scholars can examine this vast storehouse of information, posing new questions, formulating new hypotheses, and looking for new answers in the responses of participants studied generations earlier.

The verity of any research finding is strengthened when it is demonstrated across studies, across generations, across geographical regions, and across investigators. Sophisticated statistical packages and advances in meta-analyses make it possible to compare findings from several studies at once. For instance, a student working on his or her thesis may want to know whether the findings from cross-sectional studies about the significant correlation between more social support and less cognitive decline also could be confirmed by looking at these same relationships in longitudinal studies of earlier epochs. The archived raw data can be easily loaded into the student's computer for analysis of this new question.

Finally, because substantial scientific interest in cognitive aging is relatively new, much interesting work begs attention. Following are two examples. First, what are the characteristics of people who benefit most from cognitive training? What are the mediating effects of variables such as education, motivation, temperament, and degree of impairment? Do people with more education and higher levels of motivation, who are more extraverted and less intellectually impaired, benefit more from cognitive training? Or, is it just the opposite—do those people with less education and lower levels of motivation, and who are less extraverted and more intellectually impaired, make greater gains following training?

The second example is Public Law 99-592. Its passage solved one problem but exposed another. What standards will replace the arbitrary age limits of 55, 60, or 65 years to determine when someone is no longer able to do the job? Because of the currently shrinking labor pool, valid and fair procedures are urgently needed to determine whether an older person continues to have the skills necessary to perform competently in an occupation. So far, no techniques have been developed that have the necessary sensitivity to identify only those individuals with true cognitive

impairment without falsely identifying others who, in fact, do not suffer from cognitive impairment.

Knowledge about cognitive aging has accrued rapidly during the past 25 years or so. More is on the way. As these new findings accumulate, they raise the probability that all individuals, through their behaviors in small increments, will be able to positively influence a significant portion of the cognitive quality of their later lives.

See Also the Following Articles

Aging and Competency ■ Aging, Cognition, and Medication Adherence ■ Anxiety Disorders in Late Life ■ Cognitive and Behavioral Interventions for Persons with Dementia ■ Dementia in Older Adults ■ Depression in Late Life ■ Elder Caregiving ■ End of Life Issues ■ Personality and Emotion in Late Life ■ Psychoneuroimmunology ■ Psychotherapy in Older Adults

Further Reading

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as a foundation of developmental theory. *American Psychologist*, 52, 366–380.

Birren, J. E., & Schaie, K. W. (Eds.). (2001). *Handbook of the psychology of aging* (5th ed.). San Diego: Academic Press.

Cohen, S. (1996). Psychological stress, immunity, and upper respiratory infections. *Current Directions in Psychological Science*, 5, 86–90.

Craik, F. I. M., & Salthouse, T. A. (Eds.). *The handbook of aging and cognition*. Hillsdale, NJ: Lawrence Erlbaum.

Damasio, A. R. (1994). *Descartes' error: Reason, emotion, and the human brain*. New York: Avon Books.

Diamond, M. C. (1988). *Enriching heredity: The impact of the environment on the anatomy of the brain*. New York: Free Press.

Glaser, R., & Kiecolt-Glaser, J. (Eds.). (1994). *Handbook of human stress and immunity*. San Diego: Academic Press.

Heaton, R. K., Grant, I., & Damasio, A. R. (1981). Normative observations on neuropsychological test performances in old age. In I. Grant, & K. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 100–120). New York: Oxford University Press.

Kosslyn, S. M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience*. New York: Free Press.

Lexak, M. D. (1983). *Neuropsychological assessment*. New York: Oxford University Press.

McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., Ahern, F., Petrill, S. A., & Plomin, R. (1997). Substantial genetic influence on cognitive abilities in twins eighty or more years old. *Science*, 276, 1560–1563.

Poon, L. W., Rubin, D. C., & Wilson, B. A. (Eds.). (1989). *Everyday cognition in adulthood and later life*. New York: Cambridge University Press.

Rowe, J. W., & Kahn, R. L. (1987). Human aging: Usual and successful. *Science*, 237, 143–149.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum.

Salthouse, T. A. (1992). *Mechanisms of age–cognition relations in adulthood*. Hillsdale, NJ: Lawrence Erlbaum.

Cognitive and Behavioral Interventions for Persons with Dementia

Jiska Cohen-Mansfield

Research Institute on Aging, Hebrew Home of Greater Washington, Rockville, Maryland, USA

GLOSSARY

affect A person's mood or state of mind such as depression or happiness.

agitation Inappropriate verbal, vocal, or motor activity that is not judged by an outside observer to result directly from the needs or confusion of the individual; these behaviors have been labeled problem behaviors, disruptive behaviors, challenging behaviors, disturbing behaviors, behavioral problems, and agitation, all terms that are used interchangeably.

functional status—ADL The ability of a person to complete activities of daily living (ADL), that is, activities required for personal self-maintenance; some of these activities include bathing, dressing, grooming, feeding, and toileting.

self-affirmation The enhancement of the sense of selfhood/ personhood; it can be achieved through things such as meaningful activities, social interactions, and successful accomplishment of objectives.

Cognitive and behavioral interventions for persons with dementia include a large range of tools used to enhance cognitive, behavioral, and affective functioning as well as self-maintenance. Such interventions should follow a detailed analysis of the problems, goals, and abilities of these persons. The interventions are described under the following categories: cognitive, behavioral, self-affirming, sensory, social contact, structured activities, environmental, medical/nursing care, and staff/caregiver training and management. Although these interventions have shown promise in prior research, the field of investigation is still in the early stages, and larger controlled intervention studies are needed. Furthermore, research is needed to clarify the process of intervention and which intervention ingredients are crucial for success. These studies would lead the way to effectiveness and cost analyses.

1. INTRODUCTION

The scope of cognitive and behavioral interventions for persons with dementia can be conceptualized from a narrow or a wide perspective. The narrow perspective includes only interventions traditionally referred to as

cognitive behavioral, including reinforcement techniques and cognitive restructuring. A wider perspective includes a range of techniques based on different theoretical frameworks and disciplines, many of which are used either to prevent or to substitute for use of psychoactive medication. This article uses the latter approach, with the goal of clarifying the wide range of techniques available to clinicians in enhancing the quality of life of persons with dementia.

The various intervention techniques that are used with persons with dementia can be organized along several dimensions: the function of the intervention, the theoretical framework serving as a basis for the intervention, the type of activity undertaken during the intervention, and the population subgroups for whom a technique is appropriate. Interventions should be ranked by their effectiveness within a subgroup for certain goals with specific outcome criteria. Some of these dimensions are reviewed briefly in what follows.

2. PURPOSE OF INTERVENTIONS

Cognitive and behavioral interventions are used for improving functions in many domains, including the following:

- Cognitive goals include enhancement of memory as measured by memory tests, enhancement of memory for daily tasks, enhancement of ability to communicate, an increase in task engagement, an increase in orientation to time and place, and a decrease in overall cognitive decline.
- Functional goals usually involve an increase in independence or in participation in the performance of activities of daily living (ADL) (e.g., dressing, eating, toileting) and a decrease in risk of falling.
- Affective goals include a decrease in negative affect (e.g., depression, anxiety) and an increase in positive affect (e.g., pleasure, contentment).
- Behavior change involves a decrease in problem behaviors (agitation) such as aggressive behaviors, repetitive nonaggressive and restless behaviors, and repetitive verbal/vocal behaviors. Such behaviors are frequently indicative of discontent in persons with dementia and may also be disruptive to caregivers.
- In terms of self-affirmation, a loss of sense of self and of identity is a common characteristic of dementia. Even when the inner sense of self may remain intact, the socially defined self is often diminished by the decline in social, work, and leisure roles. Therefore,

some interventions target enhancement of the sense of selfhood/personhood in dementia.

3. THEORETICAL FRAMEWORKS

Many theoretical frameworks underlie cognitive and behavioral interventions. Some of the basic concepts used in these frameworks are described in what follows.

3.1. Cognitive

Research on cognitive processes differentiates between explicit and implicit memory. Explicit memory requires the ability to produce the memorized material and usually involves conscious knowledge of the material. In contrast, implicit memory affects behavior but does not involve awareness of the memory. Explicit memory is usually lost earlier in the dementia process than is implicit memory. Procedural memory involving everyday activities frequently involves implicit memory. Therefore, some simple procedures can be learned during the early and middle stages of dementia even when there is no awareness of the learning process.

3.2. Behavioral Theory

Behavioral theory asserts that behavior manifestation is affected by its antecedents and consequences. Environmental stimuli can become associated with a behavior and trigger it, whereas consequences such as reinforcement can affect the likelihood of the behavior's manifestation. To change a behavior, the contingencies between a behavior and its consequences need to be changed, for example, withdrawal of reinforcement or introduction of reinforcement for different behaviors under different circumstances. Alternatively, an intervention can alter the environmental stimulus that triggers the behavior.

3.3. Cognitive and Behavioral Theories of Depressed Affect

Cognitive and behavioral theories of depressed affect include (a) cognitive theory that ascertains that depression is a consequence of negative and distorted cognitions about oneself (based on work by Beck and Ellis), (b) theory that claims that depression involves learned helplessness and a sense of loss of control (based on Seligman's work), and (c) theory that depression relates

to an insufficient level of reinforcements or pleasurable experiences (based on Lewinsohn's work).

3.4. Unmet Needs Theory

The unmet needs theory asserts that people with dementia manifest inappropriate behavior due to unmet needs. The theory claims that people with dementia have the same needs as do other people. These needs, which are generally summarized under Maslow's hierarchy of needs, include physiological, safety, love and belonging, esteem, and self-actualization needs, with the latter sometimes being manifested as a need for cognitive stimulation or for meaningful activity. According to this theory, these needs frequently go unrecognized by caregivers, and even when they are recognized, these needs are often not satisfied because caregivers do not understand their significance or do not know how to address them. The unmet needs result in inappropriate behavior in one of several mechanisms. First, the inappropriate behavior may be a call for help due to the unmet need. Second, the inappropriate behavior may be an attempt by the person with dementia to respond to his or her need (e.g., self-stimulation as a response to sensory deprivation). Finally, the inappropriate behavior may be a direct result of the unmet need (e.g., screaming due to pain).

3.5. Humanistic Theory

Humanistic theory maintains that a person's subjective understanding of his or her experiences is more important than objective reality.

3.6. Stages of Human Development

Erikson described eight stages of human development culminating in the eighth stage (age 65 years or over), described as "Integrity versus Despair," in which a person looks back at accomplishments and ahead to the unknowns of death and feels either fulfilled or saddened. Feil added a ninth stage to Erikson's developmental framework that she termed "resolution versus vegetation." It is viewed as an opportunity to resolve conflict and achieve a sense of integrity before death.

4. TYPES OF INTERVENTIONS

This section includes a descriptive list of interventions used with people with dementia. In understanding the interventions on this list, several underlying principles need to be clarified.

4.1. General Issues in Considering Types of Interventions

• An intervention should follow a thorough assessment such as a functional analysis that examines the nature of the behavior as well as possible causes, antecedents, and consequences. The assessment should also include information about past stresses, identity, habits, and preferences as well as current medical and psychosocial status. This assessment would guide the understanding of the etiology of presented symptoms as well as realistic goals and options for treatment. Once an intervention is implemented, it should be evaluated. If it does not meet the goals, a different treatment alternative should be taken. Such alternatives might be a different intervention or might involve a different aspect of the intervention (e.g., different timing, differing amounts [dosage], different presentation style, different therapist).

• The list of interventions is not exhaustive but provides examples of the currently available interventions. Interventions are in the process of being developed and tested, so the list is an evolving body of knowledge.

• Interventions often have multiple purposes, and there are multiple options for interventions in any functional domain (see Table I). For example, cognitive interventions can be used for improving cognition, function, behavior, and/or affect.

• The descriptions in what follows provide only a general concept of the various interventions. However, each type of intervention can be considered as a whole range of possible interventions and, therefore, can be delivered in many different ways. It is the specific operational details that may make any treatment effective.

• The manner in which an intervention is presented can be crucial for its success.

• An important issue in understanding the utility of interventions is discerning the active ingredient of the interventions. Two interventions might seem quite different, yet the potent factor in both may be the opportunity to interact with a supportive person. The understanding of the underlying differences among interventions should be a focus for future research.

• Matching interventions to the person and the purpose is important. Cognitive and behavioral interventions cannot be applied uniformly to all persons with dementia. Interventions are appropriate for specific

TABLE I
Nonpharmacologic Interventions: Examples of Type of Intervention by Goal

Goal/type	Treatment modalities	Environmental
Enhance cognitive performance	Cognitive remediation/memory training Spaced retrieval Reality orientation Cognitive tasks, such as reading group activities or sorting activities	Memory books (Bourgeois et al., 2001; Brooker & Duce, 2000) Use of signs and pictures to enhance orientation
Enhance functional performance	Independence promoting strategies (Beck, 1998; Namazi, Gwinnup, & Zadorozny, 1994) Prompted (Schnelle, 1990) and scheduled voiding (Jarvis, 1981)	Increased cue visibility. Change lighting levels (Beck & Heacock, 1988; Koss & Gilmore, 1998) Increase toilet visibility, raise toilet seat
Enhance self-perception	Reminiscence therapy (Gagnon, 1997; Hoerster et al., 2001) Validation therapy (Morton & Bleathman, 1991; Toseland et al., 1997) Self-identity based interventions	Provide objects (e.g., furniture, photographs) that were meaningful in the past
Enhance pleasure	Interventions based on past or present pleasant activities Sensory interventions Structured activities	Provide objects (e.g., pictures, piano) that were liked in the past
Decrease behavior problems	Sensory interventions Behavioral interventions Social contact interventions Structured activities Staff training Environmental interventions Medical/Nursing care interventions Combination therapies (Cohen-Mansfield, 2001)	Allow access to a secure outdoor garden

Table © Cohen-Mansfield (2004).

goals and need to be matched to individuals based on their cognitive and sensory abilities as well as on their identities, habits, and preferences. In addition to tailoring the type of intervention to the person and their specific needs, the content of the intervention, such as the type of music that is played, needs to be matched to the person's preferences. This need to individualize treatments makes both research and practice more complex.

4.2. Cognitive Interventions

4.2.1. Memory Training

Positive results have been described for memory training with older persons. However, these are usually conducted with persons with either no memory impairment or mild cognitive impairment. Interventions reported for persons with dementia usually involve more support and less reliance on the person than traditional memory training methods. Interventions such as mnemonics are therefore usually inappropriate for this population. However, positive results have been demonstrated in small studies that involved daily cognitive stimulation by caregivers and with a computerized cognitive remediation system created by Butti and colleagues.

4.2.2. Cognitive Restructuring

Cognitive interventions based on Beck's and Ellis's work have been used to treat depressed persons with mild levels of dementia. These interventions involve cognitive restructuring, that is, challenging or changing distorted thought processes and thoughts that cause depressed affect.

4.2.3. Task Simplification/Cognitive Prosthesis

The most common and straightforward cognitive methods involve task breakdown, enhanced instruction, modeling, rehearsal, cueing, and gradual approximations of a task. These principles apply across different domains. The notion of simplifying tasks can also be applied across daily activities such as by providing consistency in caregiver assignment (this helps to orient older persons across different tasks). Similarly, many environmental interventions have been used to simplify orientation such as placing items that are frequently needed in a permanent location, using labels and signs, wearing a large wristwatch, and placing a card with important information in a pocket.

4.2.4. Memory Books

Memory books are booklets with autobiographical, daily schedule, and problem resolution information. When nursing assistants were trained to use the memory books, nursing home residents showed improvement in conversation and affect.

4.2.5. Spaced Retrieval for Procedural Memory

Spaced retrieval is a memory training technique that involves repeated rehearsal, with testing occurring over gradually increasing time intervals. For example, the resident may be trained to check whether his or her glasses are being worn. The resident would first be asked to check immediately after instruction, then after 20 seconds, after 40 seconds, after 1 minute, after 2 minutes, and so on. In this way, people with dementia are able to learn procedures so as to consistently check their calendars. This procedural learning is essential for independent functioning and is based on the use of procedural implicit memory, which is retained until the later stages of dementia.

4.2.6. Reality Orientation

Reality orientation is an intervention in which staff members present orienting information, including information about time, place, and person, to persons with dementia. Such information may be conveyed in special group sessions, which may include discussion of current events or may be ongoing in interactions between (usually formal) caregivers and persons with dementia.

4.2.7. Cognitive Tasks

Cognitive tasks include both group and individual activities that involve use of cognitive and memory skills. Group examples include "question-asking readings" in which a group reads a relatively low-reading level script that is typed in large font and high contrast and that is accompanied by questions typed on cards that allow participants to discuss related topics. Another group memory task is "memory bingo," in which participants have bingo-type cards with endings of popular sayings. For each card, the group facilitator reads the beginning of the saying, and participants use long-term memory to complete the sentence and use recognition skills to find it on the card in front of them. The sayings can also be used to stimulate discussion. Individual cognitive tasks include activities such as card or object sorting by category.

4.3. Behavioral Interventions

4.3.1. Differential Reinforcement

Differential reinforcement may involve positive reinforcement contingent on nonagitated behavior or "time out" (e.g., moving the person to a quiet area when agitated) or "restriction" (e.g., denying the person goods [e.g., candy], activities, or access to a location or another person when agitated).

4.3.2. Stimulus Control

Stimulus control is based on behavior being emitted under specific antecedent conditions and not others. Stimulus control involves changing the stimuli that tend to trigger a behavior, so that the behavior will be less likely to be triggered, or changing the association between the antecedent stimulus and the behavior by changing contingencies depending on the presence of various antecedent stimuli. Camouflaging exit doors and placing stop signs on them, along with training that the stop sign means stop and walk away, are examples of the use of stimulus control.

4.3.3. Maximizing Control

Helplessness has been shown to relate to depressed affect, whereas control is associated with well-being. Providing residents with opportunities to exercise control can involve having a door that they can open to an outdoor area, having a plant that they can care for, or allowing them to make decisions about meals, clothes, and the like.

4.3.4. Increasing Levels of Pleasant Activities or Noncontingent Reinforcements

Clarifying which activities or experiences are reinforcing to individuals with dementia and providing these activities or experiences is a method that has been used to treat depressed affect in persons with dementia.

4.4. Self-Affirming Interventions

4.4.1. Reminiscence Therapy

Reminiscence therapy encourages persons with dementia to recall their pasts. Sessions may use audio-visual aids such as old family photos and other objects. Reminiscence can enhance patients' sense of identity and sense of worth and/or general well-being, and may also stimulate memory processes.

4.4.2. Validation Therapy

Validation therapy, developed by Feil, involves communication with a therapist who accepts the disorientation of the person with dementia and validates his or her feelings. The assumption is that the person returns to unfinished conflicts in the past, around which a meaningful conversation can take place addressing the emotions that are important for the elderly person rather than trying to correct the disorientation. Thus, it fits within the perspective of humanistic theory. Validation therapy can take place in one-on-one treatment or in a group.

4.5. Sensory Interventions

4.5.1. Massage/Touch

Massage or therapeutic touch has been used to decrease behavior problems in people with dementia. Touch is sometimes considered a form of communication when verbal communication is no longer available. In many of the articles reporting on massage therapy, the procedures took approximately 5 minutes and were performed once or twice a day.

4.5.2. Music and White Noise

Music has been used for decreasing behavior problems and for improving cognition or function. Music has been used to relax patients during meals or bathing or as a means of providing sensory stimulation to people who are understimulated. Music interventions take many forms, including listening to a music tape, playing musical games, dancing, moving to music, and singing. Prior to using music therapy, hearing would need to be checked, and an amplifier, headphone, or hearing aid might need to be used. Music is more effective when it is individualized to match the person's preferences. White noise has also been used to induce relaxation, thereby improving sleep and decreasing restlessness.

4.5.3. Sensory Stimulation

Sensory stimulation involves the presentation of stimuli that affect different sensory modalities, including hearing, touching, seeing, and smelling. Therefore, it can include elements of the massage and music interventions described previously as well as aromatherapy, moving lights, pictures, and the like. The "Snoezelen" sensory stimulation program, which was developed in The Netherlands and includes a variety of relaxing stimuli, is also a type of sensory stimulation.

4.6. Social Contact Interventions

Loneliness is highly correlated with depressed affect, and being alone has been shown to relate to behavioral problems. The best intervention for loneliness is positive interaction with a person who is meaningful to the elderly person. However, this is often not feasible. Therefore, a variety of alternative social interventions have been developed. In addition to the following social contact interventions, most group activities, including cognitive ones, can be used as vehicles to promote social contacts.

4.6.1. One-on-One Interaction

One-on-one interaction is a potent intervention for loneliness and behavior problems. It can be conducted with relatives, paid caregivers, or volunteers.

4.6.2. Pets/Dolls

Pet therapy often involves visits with a dog that last from a half-hour to an hour. However, other pets (e.g., a cat, fish) can be used as well. Even plush stuffed animals have been used successfully, as have robotic pets. Dolls have also been used to simulate companions, since they can be viewed as real babies. Pet therapy not only involves interaction with the pet but also serves as a topic for interaction with other people.

4.6.3. Stimulated Interaction/Family Videos/ Interaction Videos

Simulated presence therapy uses an audiotape to simulate phone interactions. The tape contains a relative's portion of a telephone conversation and leaves pauses that allow the older person to respond to the relative's questions. Family videotapes also simulate interaction by having a loved one talking to the person with dementia. These videotapes have been found to produce engagement and decrease agitation in persons with dementia. Interaction videos are those in which the persons on the videos, usually professionals, interact with the persons with dementia, often recalling information from the past or inviting the persons with dementia to sing along, using long-term memory for well-known songs.

4.7. Structured Activities

4.7.1. Indoor Activities

Structured activities are used to improve affect and decrease behavior problems. They can take many forms and include activities that Buettner termed "simple pleasures," including both group and individual activities. The activities may involve manipulation (e.g., ball throwing), nurturing (e.g., watering a plant), sorting, cooking, sewing, or engaging in sensory intervention as described previously (e.g., music, tactile stimulation with a fabric book). The content of activities may be based on information regarding which activities were or are reinforcing to the individual.

4.7.2. Outdoor Activities

Outdoor walks and physical activities are forms of structured activities that have the potential to improve function and affect and to decrease problem behaviors. Outdoor walks may take place in the company of a caregiver and so may also involve a social component, or they may occur in a secure outdoor area involving an environmental intervention. Outdoor walks are often limited by weather conditions.

4.8. Adapted/Enhanced Environmental Interventions

Environmental interventions can be used for multiple goals such as to decrease or accommodate behavior problems, to enhance functional status, and to increase orientation.

Environments for increasing functional ability have included different levels of environmental adaptation, from decreasing clutter to providing grab bars or handrails. The following are specific types of adaptation or enhancements.

4.8.1. Visibility of Cues Needed for Activities of Daily Living

The importance of visibility has been demonstrated in several research studies. For example, use of toilets increased when toilets were more visible, and eating behavior improved with better light and increased contrast between plates and the table.

4.8.2. Seating in Motion

Rocking chairs and gliding swings have been used for relaxation, improved affect, stimulation, and reduced physical agitation.

4.8.3. Homelike Ambience

Homelike furnishings in the institutional setting and enhanced ambience can affect many aspects of well-being for both patients and their caregivers.

4.8.4. Reduced Falls

Decreasing bed height, placing mattresses on the floor, using hip protectors, improving light on the way to the bathroom, and improving call systems all can help to decrease the risk of injury due to falling.

4.8.5. Privacy and Intrusion Deterrence

To prevent trespassing into other people's rooms or through emergency exit doors, doors and doorknobs can be camouflaged with cloth panels or murals, thereby disguising the doors. In addition, providing alternative doors, which can be controlled by the patient and permit movement into another secured area, can be useful in reducing trespassing.

4.8.6. Clear Labels and Signs

Clearly labeled signs with large font and high contrast can be used to increase orientation. Nursing homes can also use boxes with pictures of residents taken in

the past, or other memorabilia that is meaningful to them, to help orient residents in finding their rooms. Clear labels on public toilets are useful in decreasing incontinence.

4.8.7. Wandering Areas

Places that allow people to walk in a safe environment can be created either indoors or outdoors. These usually involve some type of walking loop that allows for walking but does not have an exit door that might allow for egress into an insecure area. In contrast, exit doors that can be opened by residents and allow them control of their own exit to another secure area can be therapeutic.

4.8.8. Peace and Quiet

Reduced stimulation environments, such as quiet rooms with soft colors, an absence of paging systems, and only a few objects, have been reported to be helpful. Reduced stimulation interventions may use neutral colors on pictures and walls and a consistent daily routine, with no televisions, radios, or telephones (except one phone for emergencies).

4.8.9. Nature Environments

Natural environments, including pictures of fish and/or sounds of bird songs or bubbling brooks, have been used to decrease agitation in the shower. Visual, auditory, and olfactory stimuli have been used to make nursing home corridors feel like home or like a natural outdoor environment, thereby improving resident and staff well-being and decreasing trespassing and exit-seeking behaviors.

4.9. Medical/Nursing Care Interventions

Although medical and nursing care interventions would not usually be considered cognitive or behavioral treatments, they are included here because it is the behavioral assessment that is used to reveal the underlying etiology of negative affect or behavior problems that allows a medical or nursing intervention to relieve the condition. Therefore, interventions such as light therapy to improve sleep, pain management, reduction of discomfort by improved seating or positioning, and removal of physical restraints all have

been associated with improvement in behavior. Similarly, the provision of eyeglasses or hearing aids can be an important intervention for treating sensory deprivation and loneliness as well as the ensuing depressed affect and problem behaviors.

A behavioral approach may also guide nursing interventions directly in helping elderly persons to achieve maximal functional levels. One example is that of cognitive interventions that include cueing and task breakdown. Another example involves toileting management protocols. Two types of toileting protocols have been described. In scheduled or timed voiding toileting, patients are taken to the toilet either at fixed times (usually every 2 hours) or on a schedule that is based on their voiding pattern. Prompted voiding involves asking residents on a regular basis whether they need assistance with toileting. Patients are helped when they indicate such a need. Both types of protocols can be effective in reducing incontinence.

4.10. Staff/Caregiver Training and Management

Staff training programs can take many forms and focus on any of the interventions outlined previously. Many focus on improved understanding of older persons and the impact of the disease and on improving verbal and nonverbal communications with persons suffering from dementia. Given that good communication skills are crucial for proper assessment of a problem area and for delivery of any intervention, training for such communication is the basis of good caregiving. Examples of methods for simplifying and clarifying communication using both verbal and non-verbal communication channels were provided by Beck and Heacock in 1988. Changing caregiver behavior through training is a complex and difficult challenge and often requires ongoing instruction, modeling, monitoring, feedback, and support of the caregiver. Therefore, in institutional settings, staff training is closely tied to management.

4.11. Combination Therapies

Good care involves using combinations of the available interventions and tailoring them to the needs, abilities, and preferences of the older person. In addition, many intervention programs combine elements from various intervention modalities.

5. EFFICACY, EFFECTIVENESS, AND UTILITY

There are two basic questions concerning interventions. First, are interventions effective? Second, what is the cost of interventions? There is insufficient research to answer either question, although partial answers have been suggested. One study reported the benefits of validation therapy for reducing aggressive behavior, and another reported on the benefits of reality orientation on orientation and affect, in comparison with control groups. However, in a review of reality orientation, validation therapy, and reminiscence therapy, Gagnon concluded that reality orientation and validity therapy do not produce sufficient change to justify their costs. Cohen-Mansfield reviewed 83 studies of nonpharmacological interventions for inappropriate behaviors in dementia and described the majority as reporting a positive, but not always significant, impact. Many of the studies included small samples and other methodological limitations, most often resulting from limited funding for this type of research.

There are few studies that have examined the question of cost. One study by Rovner and colleagues implemented an intervention that included daily group-structured activities, reduction of psychotropic medications, and weekly educational rounds with staff members. They calculated the cost of the 6-month intervention program at $8.94 per patient per day. Another study by Schnelle *et al.* examined ways in which to improve nursing practices concerning exercise and incontinence care and concluded that current staffing levels are inadequate to provide good care.

Given the early stage of research in this field, there is an urgent need to increase our understanding about the following basic questions. Which interventions are efficacious for which individuals? Which aspects of an intervention are necessary for it to be efficacious? What are the active ingredients, or principles at work, in various interventions? Which personal characteristics (e.g., gender, culture, prior stress) need to be considered when matching an intervention with a person? Process variables (e.g., the person who is delivering the intervention, the timing of the intervention) and their significance need to be elucidated. Only after we have answers to these basic questions can the questions of effectiveness and costs be addressed properly.

See Also the Following Articles

Aging and Competency ■ Cognitive Aging ■ Cognitive Skills: Training, Maintenance, and Daily Usage ■ Dementia in Older Adults ■ Elder Caregiving ■ End of Life Issues ■ Personality and Emotion in Late Life ■ Psychotherapy in Older Adults

Further Reading

Beck, A. (1997). The past and future of cognitive therapy. *Journal of Psychotherapy Practice & Research*, 6, 276–284.

Beck, C., & Heacock, P. (1988). Nursing interventions for patients with Alzheimer's disease. *Nursing Clinics of North America*, 23, 95–124.

Camp, C. J., Foss, J. W., O'Hanlon, A. M., & Stevens, A. B. (1996). Memory interventions for persons with dementia. *Applied Cognitive Psychology*, 10, 193–210.

Cohen-Mansfield, J. (2000). Theoretical frameworks for behavioral problems in dementia. *Alzheimer's Care Quarterly*, 1(4), 8–21.

Cohen-Mansfield, J. (2001). Nonpharmacologic interventions for inappropriate behaviors in dementia: A review, summary, and critique. *American Journal of Geriatric Psychiatry*, 9, 361–381.

Day, K., & Calkins, M. P. (2002). Design and dementia. In R. B. Betchel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 374–393). New York: John Wiley.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Feil, N. (1982). *Validation: The Feil method.* Cleveland, OH: Feil Productions.

Gibbons, A. C. (1988). A review of literature for music development/education and music therapy with the elderly. *Music Therapy Perspectives*, 5, 33–40.

Lewinson, P. M., & Youngren, M. A. (1976). The symptoms of depression. *Comprehensive Therapy*, 2(8), 62–69.

Opie, J., Rosewarne, R., & O'Connor, D. (1999). The efficacy of psychosocial approaches to behaviour disorders in dementia: A systematic literature review. *Australian and New Zealand Journal of Psychiatry*, 33, 789–799.

Schnelle, J. (1990). Treatment of urinary incontinence in nursing home patients by prompted voiding. *Journal of the American Geriatrics Society*, 38, 356–360.

Seligman, M. E. P. (1974). Depression and learned helplessness. In R. J. Friedman, & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research.* Washington, DC: Winston–Wiley.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death.* San Francisco: Freeman.

Cognitive Skills: Training, Maintenance, and Daily Usage

Karlene K. Ball, Virginia G. Wadley, David E. Vance, and Jerri D. Edwards

University of Alabama, Birmingham, Alabama, USA

GLOSSARY

divided attention The ability to split a person's attentional focus or to maintain vigilance with respect to two or more objects or activities.

executive function A set of higher order cognitive abilities primarily associated with frontal and prefrontal structures of the brain; it involves planning, organizes information, inhibits responses, and orchestrates mental resources.

hippocampus A subcortical structure of the brain that is key for memory function and is specifically involved in encoding new information, a vital first step for subsequent recall.

mild cognitive impairment (MCI) A term used to describe individuals with focal memory impairments but no evidence of frank impairment in other cognitive domains and no evidence of impaired daily functioning; MCI is often thought of as a preclinical stage of dementia.

transfer of training The capacity for training-induced improvements in cognitive skills to translate into improved performance of cognitively demanding activities.

Cognitive skills are basic and higher order mental abilities, such as attention, information processing, memory, executive functioning, and reasoning, that interact with sensory and perceptual systems to determine success in performing daily activities. Older adults are at risk for cognitive decline but may benefit from strategies to maintain cognitive skills and from training programs to enhance certain skills.

1. COGNITIVE SKILLS AND THEIR IMPORTANCE FOR DAILY FUNCTIONING

The ability to think, remember, and reason allows people to negotiate the world in which they live, whether it be planning the steps in cooking a meal, remembering to take their medications on time, deciding how to spend their money, or reacting quickly to avoid hitting a dog in the road. People's cognitive abilities provide them with the flexibility to engage in a number of unique and demanding situations. Declines in any of these abilities can subsequently diminish the ability to function in day-to-day life. Cognitive abilities that tend to decline with age, with potential impact on the ability to perform everyday tasks, include attention, speed of processing, memory, executive functioning, and reasoning. These abilities interact with sensory and

perceptual systems to determine performance in daily life.

1.1. Sensation and Perception

Information processing must begin with acquiring the information in the first place. Thus, age-related changes in sensory–perceptual abilities (e.g., poorer vision, hearing) can influence older adults' ability to perform measures of cognitive ability as well as everyday activities. For the most part, older adults can compensate for age-related sensory decline through corrective lenses and /or hearing aids. Older adults also compensate for degraded sensory information through cognitive strategies in which they use the context of a situation (e.g., conversation, other words in a sentence, situational cues) to understand what is being presented. However, impairments in sensory processing can lead not only to poorer performance on cognitive tests, and perhaps a false impression of cognitive impairment, but also to restrictions in everyday activities such as driving. Such restrictions may, in turn, lead to increased depression and/or the need for formal care.

1.2. Attention

Attention encompasses a rather broad range of abilities, from simply orienting (e.g., turning attention toward a sudden noise), to sustained attention or vigilance (e.g., maintaining focus over an extended period of time). Divided attention is quite commonly used in everyday contexts and occurs when people try to do two things at the same time. This ability may or may not decline with age, depending on the individual and the difficulty of the two tasks being performed at the same time. The use of cell phones while driving is an example of a divided attention task that can be fairly difficult. There is mounting evidence that crash risk is elevated for drivers of all ages while using cell phones, and this risk may increase with age. In general, research indicates that older drivers are susceptible to decrements in driving performance when increased demands are placed on attention, whether it be cell phones or multiple events occurring at a congested intersection. Selective attention, or the ability to pay attention to relevant information while ignoring irrelevant information, is also germane in this context. It has generally been found that the detrimental effect of irrelevant information, or distraction, is higher for older adults.

1.3. Speed of Processing

Speed of processing, or the ability to perceive and process information quickly, is one of the first cognitive abilities to decline with age. Over the past decade or so, many studies have demonstrated relationships between cognitive processing speed and everyday function in older adults. Various indexes of mobility have been linked to processing speed impairments in older adults with and without dementia. For example, reductions in life space, or the extent of a person's travel in his or her environment inside and outside the home, have been associated with impairments in processing speed. Furthermore, multiple studies have demonstrated that processing speed impairment is an excellent predictor of increased crash involvement in older drivers and is an even stronger predictor of injurious crashes. Reduced processing speed also is related to slower performance of instrumental activities of daily living (IADLs) such as looking up a telephone number, finding items on a crowded shelf, and reading medication labels.

1.4. Memory

Memory is one of people's most obvious cognitive abilities. Of all the mental abilities, people report complaints about this ability more often than about any other ability, probably because it is fundamental to most everyday tasks. Also, memory decline can be obvious when a person is attempting to recall necessary information such as telling a doctor what medications he or she is taking or trying to remember a phone number. Although activation of both cortical and subcortical areas of the brain is necessary for good memory functioning, the hippocampus, a subcortical structure, has the primary responsibility of encoding new information, a vital first step for subsequent recall of the information. As seen with Alzheimer's disease, when the hippocampus is severely damaged, profound memory loss is experienced. Although some normal age-related decline in memory does occur with age, such changes generally do not severely affect daily functioning to the same degree as found in a dementia such as Alzheimer's disease. Regardless, age-related memory declines have been found to influence activities of daily living (ADLs), whether they be forgetting to adhere to medical appointments or not remembering to pay bills. Research has linked memory declines to self-reported impairment with daily shopping, preparing a hot meal, housecleaning, managing finances, and completing forms.

1.5. Executive Function

One of the most complex cognitive abilities is executive function, which includes people's ability to plan, organize information, inhibit responses, think abstractly, and reallocate mental resources. This ability is primarily associated with the frontal lobes of the brain and is known to decline with age. Executive functioning is necessary to complete tasks that require complex behavior or have multiple steps. For instance, executive functioning is useful in paying bills because this task requires planning (e.g., determining how much money to transfer from a savings account to a checking account), calculating how much money remains in the checking account after paying the bills, and (in many cases) deciding what is the least amount that can be paid to prioritize money to other uses. Thus, financial capacity can be impaired in individuals with executive dysfunction, as is seen in adults with Alzheimer's disease. However, dysfunction in executive abilities can also affect less cognitively demanding activities. For example, scientists have found that lower levels of executive functioning are related to impairment in the ability to perform other IADLs such as medication management, even among noninstitutionalized retirees.

1.6. Reasoning

Reasoning is another complex cognitive ability affected by aging. This ability is similar to executive functioning; in fact, damage to the fontal lobes also impairs the ability to reason. Unlike executive functioning, reasoning focuses on using logical constructs, knowledge, and principles to extrapolate a solution to a problem. Reasoning is a sophisticated problem-solving skill that requires both memory and executive functioning. Reasoning is used in a variety of real-world tasks, ranging from medical decisions to driving behavior. Thus, declines in reasoning ability can impair decision making in a variety of real-world scenarios. As with other cognitive skills, reasoning ability has been found to predict IADL performance. Scientists have found that inductive reasoning is important for figuring out problems associated with everyday tasks such as shopping and managing money. In fact, researchers have found that as inductive reasoning is enhanced, everyday problem solving also improves.

1.7. Cognitive Impairment

Whereas age-related declines in cognition and function may be relatively subtle and selective in normal aging, the declines experienced by individuals with mild cognitive impairment (MCI) or dementia are more pronounced and become pervasive over time.

MCI is a term used to describe individuals with focal memory impairments but no evidence of frank impairment in other cognitive domains and no evidence of impaired daily functioning. MCI is often thought of as a preclinical stage of dementia. Approximately 5 to 15% of adults age 65 years or over may be affected by MCI. Individuals with MCI, by virtue of their memory impairments and their risk of progression to dementia, are at high risk for declines in the abilities needed to function effectively and independently in society. These individuals are likely to experience a gradual loss of functional skills of a magnitude that lies somewhere between the subtle decrements associated with normal aging and the much more obvious deficits associated with dementia. Little has been done to directly evaluate the impact of cognitive changes in MCI on the performance of everyday tasks. A critical question is how changes in memory and other cognitive abilities influence the performance of everyday tasks over the period of time that, for many individuals, will mark the transition from MCI to dementia. To date, no one has mapped cognitively demanding everyday activities to component cognitive abilities in MCI. Thus, it is difficult to determine whether some tasks of daily life are dependent on specific abilities or how levels of ability influence functional performance. A clearer understanding of this relationship within the MCI population is needed.

Dementia includes a category of diseases, such as Alzheimer's disease, characterized by more severe cognitive impairments that hamper a person's ability to perform everyday tasks. The incidence of dementia rises with age. For example, only approximately 5% of adults over 65 years of age have Alzheimer's disease, whereas 10 to 20% of adults over 80 years of age are affected, and estimates have indicated that as many as 40% of adults age 85 years or over suffer from the disease. Research in the field of cognitive aging has demonstrated that the deterioration of cognitive function that occurs with dementia contributes to a decline in the performance of everyday activities. Research has also shown a clear link between declining everyday function (e.g., managing finances, grocery shopping, driving) and subsequent hospitalizations, need for care, and death.

2. HOW COGNITIVE SKILLS MAY BE MAINTAINED

Cognitive decline in numerous areas can result in difficulty in performing tasks important for everyday life and

for maintaining independence throughout old age. Thus, an important question to consider is, "How can individuals maintain cognitive abilities and avoid cognitive decline?" Several factors have been related to sustained cognitive function with age, including education level, physical activity level, pulmonary health, and feelings of self-efficacy. By and large, healthy and active lifestyles that include proper diet and exercise, prevention of disease, and the avoidance of trauma all are important for sustaining cognitive capacity with advancing age.

Higher education levels are associated with better cognitive performance during older adulthood. Education level affects cognitive function both directly, through enhancing brain function, and indirectly, through individuals with more education being more likely to participate in intellectually stimulating activities that, in turn, preserve cognitive abilities.

Physical activity has various benefits for older adults and positively affects cognitive function in a variety of ways. For example, physical activity has been shown to decrease depression (which can cause cognitive deficits, particularly for older adults) and to increase overall physical health and improve quality of life. Strenuous exercise directly benefits brain and central nervous system functioning, which are vital to maintaining cognitive abilities. Accordingly, aerobically active individuals perform better on measures of working memory, reasoning, and speed of processing than do those who exhibit more sedentary behavior.

In addition to exercise, many facets of physical health have been related to cognitive function. For example, a specific physiological measure, pulmonary peak expiratory flow rate, has been found to be predictive of sustained cognitive ability with age. This physiological index is related to cardiovascular health and physical activity level. Studies have noted in particular that preserved vascular and cardiovascular health is vital for cognitive function. For example, decreased cerebral blood flow is associated with poor cognitive performance. Furthermore, poor cardiovascular health is associated with increased incidence of stroke, a leading cause of cognitive impairment. In the Seattle Longitudinal Study (SLS), absence of cardiovascular disease was linked to better reasoning ability and speed of processing performance for older adults. A number of other diseases, such as diabetes, emphysema, and Parkinson's disease, can also result in cognitive deficits. In general, the fewer chronic diseases a person has, the better his or her chances of maintaining cognitive abilities with advancing age.

Overall, a healthy lifestyle is important not only for physical health but also for mental health. Obesity has been related to higher incidence of many chronic diseases, which in turn are related to cognitive decline. Cigarette smoking and excessive alcohol consumption both have been linked to higher incidence of stroke, a leading cause of cognitive impairment, as well as to many other poor health outcomes, thereby negatively affecting cognition. Furthermore, alcoholism can be directly detrimental to cognitive function. Most obviously, traumatic head injury results in cognitive impairment and is associated with increased incidence of dementia later in life. Therefore, simple preventive behavior, such as using seat belts combined with air bags while in automobiles and using helmets while riding on motorcycles, bicycles, or scooters, is advised.

In addition to physical health and healthy lifestyles, psychological well-being is important for preserving cognitive function. An individual's attitude can affect his or her cognitive performance with age. For example, self-efficacy, a positive evaluation of one's own cognitive capabilities, is related to better cognitive performance. Individuals in the SLS who rated their personalities as "flexible" during midlife were found to be at reduced risk for cognitive decline in later life. Furthermore, individuals who indicated satisfaction with their life accomplishments during middle age performed better on cognitive measures in later life.

Sustained social activity also is associated with better cognitive performance. For example, researchers have found that those who engage in social, domestic, and leisure pursuits are less likely to have cognitive impairment. Similarly, the SLS indicated that older adults who had high socioeconomic status and were socially active experienced the least cognitive decline over a 7- to 14-year period. Risk of cognitive decline was lessened for individuals who read extensively, traveled, attended cultural events, were involved in clubs and professional associations, and pursued continuing education opportunities. Extensive social networks and support systems, including intact families, have also been related to better cognitive functioning with age. Conversely, depression has been identified as a psychological/biological cause of decreased cognitive function. Particularly for older adults, depression can manifest itself as cognitive impairment.

3. HOW COGNITIVE SKILLS CAN BE TRAINED AND THE IMPACT OF TRAINING

3.1. Cognitive Training

Given that cognitive abilities can and do decline with age, often resulting in difficulty in performing everyday

tasks, the possibility of cognitive training to prevent, slow, or reverse age-related cognitive decline has been investigated. A growing number of studies now support the protective effects of intellectual stimulation on cognitive abilities for older adults without dementia. Early studies in the area of cognitive training were conducted within the Adult Development and Enrichment Project (ADEPT) and the SLS. Both of these studies provided 5 hours of strategy training, preceded and followed by cognitive assessment. Significant cognitive training gains were observed in both studies for the specific abilities that were trained.

A large randomized clinical trial, Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE), recently evaluated the impact of three promising cognitive interventions—speed of processing training, memory training, and reasoning training—on the maintenance of both cognitive and day-to-day abilities in community-living older adults. The study showed that for all three interventions, there were significant and specific improvements in cognitive ability as well as an increased benefit of additional booster training. The amount of training gain for cognitive abilities was equal to or greater than the amount of decline that would be expected in older adults without dementia over 4 to 14 years of aging in the absence of any training.

3.2. Durability of Training

In the ACTIVE study, among those individuals who experienced immediate cognitive improvements, training gains were found to persist for at least a 2-year follow-up period for all three interventions. Prior cognitive training studies have also pointed to the durability of training effects, with maintenance of reasoning training demonstrated at even 7 years after original training. Those older adults in the SLS who received reasoning booster training were found to substantially outperform their initial baseline levels 7 years later. Although studies of memory training have consistently shown that memory improvement can be maintained 1 week to 6 months, results on the durability of training beyond 6 months are less consistent. Some investigators have reported a significant decrease in memory performance and a substantial decrease in mnemonic strategy use over a 3-year interval in a memory training group. Multifactorial memory training programs, such as those incorporating psychomotor training in addition to memory training, have resulted in maintenance of memory improvement extending to 3½ years.

Maintenance of speed of processing training effects has not been evaluated extensively beyond the ACTIVE study, which demonstrated maintenance for at least 2 years. In an early training study, speed training effects were found to persist over a 6-month period. Furthermore, participants in the older age group (>60 years) performed at the same level following training as the middle age group (40–59 years) had performed at baseline. In a subsequent study among older adults, training effects were found to persist over 18 months and remained significantly better than baseline. Thus, in general, cognitive training studies have supported the durability of training effects.

3.3. Transfer of Training

Given that cognitive training can result in improvements of targeted cognitive skills, can these cognitive improvements transfer to improvements in performance of everyday activities? Older adults are constantly faced with learning new things such as skills for using computers, home medical devices, automated teller machines, videocassette recorders, and job-related skills. In addition, older adults would benefit from preservation of functional abilities that were most likely acquired at a younger age (e.g., driving, other IADLs). Although the onset of decline occurs for many cognitive abilities during a person's 60s, the onset of decline for everyday abilities typically occurs during the 70s or later. Obviously, functional competence is multidimensional, relying on multiple cognitive abilities. Consequently, assessment of functional competence/everyday abilities has included multiple methods, including global measures, performance-based measures, and self-report measures.

An example of a cognitively demanding task of daily living is the Everyday Problems Test (EPT), which assesses a person's ability to interact with 14 everyday situations (e.g., doubling a recipe, interpreting transportation schedules, understanding a Medicare benefit chart). Similarly, the Observed Tasks of Daily Living (OTDL) assessment presents tasks in three domains: medication use, financial management, and telephone use. Participants are asked to perform actions required to solve everyday tasks (e.g., searching medication label for side effects, making change). Performance on these measures has been related to multiple basic cognitive abilities; therefore, it may improve with cognitive training.

Other measures of everyday ability emphasize the speed of responding to real-world stimuli. For

example, slower processing abilities can present a risk with respect to safe driving. The speed of processing research has examined age differences in the Useful Field of View (a collection of speeded tasks evaluated in the ACTIVE clinical trial) as well as the benefits of speed of processing training to both cognitive abilities and everyday tasks. Prior training research found that improved speed of processing resulted in improved braking times in a driving simulator in addition to significantly reducing the number of dangerous maneuvers made by drivers during an open-road driving evaluation and maintaining the extent of people's driving over an 18-month period. Self-reported difficulty in driving, driving avoidance, and reducing the extent of people's driving all have been found to be associated with slower processing speed. Despite the fact that individuals with impairments appear to self-regulate their driving performance by reducing their exposure to challenging driving situations, studies have shown that these modifications have not been sufficient to eliminate their increased risk of crash involvement. Another measure, the Timed IADL (TIADL) test, provides an index of the time needed to successfully complete five tasks (e.g., finding a number in a telephone book, finding specific items of food on a grocery shelf). The time needed to complete the TIADL tasks is similarly reduced following speed of processing training among older individuals with impaired processing speed. Thus, speed of processing training has been found to generalize to important everyday abilities in addition to boosting cognitive ability.

4. CONCLUSION

Clearly, there are many aspects of cognition that can potentially decline as a part of the aging process, and some of these declines can affect performance of everyday tasks. There is a growing body of literature demonstrating that these declines can be prevented or slowed through changes in lifestyle. Progress also has been made in the area of cognitive training, indicating that cognitive decline is not necessarily irreversible during older age and that interventions may allow longer maintenance of cognitive function into older age. Furthermore, the fact that many cognitive declines are amenable to moderating effects implies that the independence and mobility of individuals can be extended into later life. In short, the outlook for preventing, delaying, or reversing the onset of age-related cognitive decline is bright.

See Also the Following Articles

Cognition and Culture ■ Cognitive Aging ■ Cognitive and Behavioral Interventions for Persons with Dementia ■ Dementia in Older Adults

Further Reading

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., Morris, J. N., Rebok, G. W., Smith, D. M., Tennstedt, S. L., Unverzagt, F. W., & Willis, S. L., for the ACTIVE Study Group. (2002). Effect of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association*, *288*, 2271–2281.

Birren, J. E., & Schaie, K. W. (Eds.). (1995). *Handbook of the psychology of aging* (4th ed.). San Diego: Academic Press.

Fisk, A. D., & Rogers, W. A. (Eds.). (1996). *Handbook of human factors and the older adult.* San Diego: Academic Press.

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle Longitudinal Study.* New York: Cambridge University Press.

Willis, S. L. (1987). Cognitive training and everyday competence. In K. W. Schaie (Ed.), *Annual review of gerontology and geriatrics,* (Vol. 7, pp. 159–188). New York: Springer.

Cognitivism

Pío Tudela

University of Granada, Granada, Spain

GLOSSARY

adaptive control of thought (ACT) An influential unified theory of cognition based on a symbolic architecture of the mind, developed by John Anderson. This theory had several different versions; one of the most influential is known as ACT[*].

automatic and controlled processes A cognitive process is said to be automatic when its functioning may be triggered by the mere presence of a stimulus, without need for conscious attention. It is called a controlled process when its functioning depends upon conscious attention.

bottom-up and top-down processing Cognitive processing takes place in a bottom-up manner when the direction of the information processing flow goes from simple (or low level) to complex (or higher level) processes. Top-down processing goes in the opposite direction and refers to the influence of high level on lower level processes.

cognitive neuroscience An interdisciplinary enterprise in which cognitive psychology, neuroscience, and computer science cooperate in order to study and understand the relationship between mind and brain; considered a subdivision of cognitive science.

cognitive science An interdisciplinary enterprise in which different branches of scientific knowledge dedicated to

the study of cognition merge. Six disciplines are involved: psychology, computer science, linguistics, philosophy, anthropology, and neuroscience.

connectionism An approach within cognitive science contending that information processing takes place through the interactions of a large number of simple processing elements called units, each sending excitatory and inhibitory signals to other units. Parallel processing and distributed representations are the main characteristics of connectionist (also called PDP) models.

information processing An approach to the study of cognition that considers the human mind as a symbol-manipulating system similar to a computer. Cognition can be divided into a series of stages; at each stage, certain particular operations are performed on incoming information. The eventual performance of the system is thought to be the outcome of this series of stages and operations.

levels of explanation Different viewpoints from where the behavior of a complex information processing system may be considered, described, and eventually explained.

local and distributed representations A representation is said to be local if there is a one-to-one correspondence between entities in the represented world and elements in the representation. In a distributed representation, each entity is represented by a pattern of activity distributed over many elements in the representation, and each element is involved in representing many different entities of the represented world.

neural networks Sets of elementary, neuron-like information-processing units related to each other by excitatory or inhibitory connections. Neural networks form the basic structure of connectionist architectures.

physical symbol system A system consisting of physical symbols. A physical symbol is a physical entity able to

designate or refer to another entity that is the *meaning* of the symbol. Physical symbols form the basic structure of symbolic architectures.

serial and parallel processing Information processing is organized in a serial way when processes are carried out one after another, each taking as input the output of the previous one and beginning operation when the preceding process has finished. When processes are organized in parallel, they operate simultaneously and their input is independent of the output of other processes.

situated cognition A nonrepresentational approach within cognitive science contending that mind, body, and environment are coupled systems, dynamically interacting with each other. Hence, they should not be studied as isolated systems; rather, their mutual interactions should be the main goal of cognitive science.

Soar A general cognitive symbolic architecture for developing systems that exhibit intelligent behavior. It was created by John Laird, Paul Rosenbloom, and Allen Newell.

In the broad sense of the term, cognitivism encompasses all those theoretical approaches that consider the study of the human mind and its constituent processes as the main objective of psychological science. In this sense, psychology, as an experimental enterprise, was cognitive from its very beginning. For example, W. Wundt considered the task of psychology to be the analysis of conscious processes into its constituent elements, and many chapters in William James' book *The Principles of Psychology* were dedicated to analyze processes such as attention, memory, and consciousness, the core cognitive processes. In a more strict sense, cognitivism refers to a particular theoretical approach to studying cognitive processes that originated as a reaction against behaviorism in the midst of the 20th century and is known as the information-processing approach. This article primarily concentrates on the origin and subsequent development of this approach, though attention is also paid to other cognitive views influencing psychological research.

1. COGNITIVISM DURING THE BEHAVIORIST ERA

From a cognitive point of view, the contribution of behaviorism to our knowledge of psychological processes was almost restricted to the realms of learning, motivation, and emotion. Little if any concern was shown about cognitive processes such as perception, attention, memory, and consciousness in the writings of Hull or Skinner. Even when they made a foray into the territories of thought (C. Hull) or language (B. Skinner), their nonrepresentational approach turned out to be too restricted and had little influence on subsequent research. Among behaviorists, E. Tolman was one of the few who did not show disdain for using terms and concepts with a cognitive flavor. Even in his case, however, those terms did not refer to processes to be studied for their own interest; rather, they were conceptualized as intervening variables useful to predict behavior.

It should be mentioned, however, that behaviorism did provide a lasting contribution to cognitive and psychological research in general. Emphasis on the measurement of variables, the careful design of experiments, and the wide-ranging use of objective scientific methods is part and parcel of the invaluable legacy of behaviorism. Thanks to this methodological heritage, modern cognitivism could establish itself on a solid scientific foothold.

Despite the general sway of behaviorism during the second quarter of the 20th century in the United States, interest in cognitive research remained active within restricted domains of psychological inquiry. This was particularly so in Europe, where the influence of behaviorism took longer to be felt and was never as strong as it was in the United States. The following streams of investigation deserve to be pointed out:

1. The contribution of Gestalt psychologists to the development of perceptual theory.

2. Research on concept formation, problem solving, and thinking by researchers such as Edouard Claparède, Wolfgang Köhler, and Frederic Bartlett.

3. The sociohistorical approach to cognitive development carried out in Russia by Lev Semenovich Vygotsky.

4. Jean Piaget's research program on cognitive development and genetic epistemology in Switzerland. Piaget was one of the first thinkers to realize the interdisciplinary nature of cognitive research. In 1955 he founded the International Center for Genetic Epistemology in Geneva, where scientists from different disciplines tried to set up the foundations of a true cognitive science. Piaget influence on cognitive development and educational psychology is still noticeable today; research along his line of thinking is an integral part of contemporary cognitive research.

5. It is important to recognize that, in some areas of psychology, behaviorism never played the dominant role it did within experimental psychology. References

to mind and mentalistic concepts never disappeared from social and clinical psychology. Authors such as Kurt Lewin, Leon Festinger, Erich Fromm, Carl Rogers, and Abraham Maslow are only a few among the many who did not adhere to the dominant behavioristic approach.

All these streams of investigation paved the way for what has come to be known as the cognitive revolution.

2. THE INFORMATION-PROCESSING APPROACH

Cognitivism, as it was articulated within the information-processing approach, was from its beginning the result of multiple lines of thought coming from different scientific disciplines. Within psychology, there was a profound dissatisfaction with behaviorism as a method of tackling complex human behavior. World War II compelled psychologists to face practical problems, such as the manipulation of aircraft, radar, and other highly technical devices involving multiple stimuli and responses and posing heavy demands on human operators. In these situations, concepts from the animal learning labs had little application, so psychologists turned to communication engineering and information theory for help. The human operator was conceptualized as an information transmitter and decision maker, and research on selective attention followed as a matter of course.

In 1959, the linguist Noam Chomsky published a devastating review of Skinner's book *Verbal Behavior*; his review was also an attack on other behavioristic accounts of language. As an alternative to studying linguistic performance, Chomsky emphasized linguistic competence, knowledge about the rules of grammar that the native speaker of a language owns. Under Chomsky's influence, psychologists became interested in the cognitive processes responsible for the structural aspects of human language. Subsequently, the significance of linguistic structures for our understanding of memory and thought was increasingly acknowledged.

The most determining influence, however, came from the realm of computer science. Computers are general purpose machines that can be programmed to perform any well-defined task. As information-processing systems, they are able to receive, store, transform, and retrieve information. A computer performance may be analyzed at the physical level of circuits used to implement the system (hardware), but there is a different, more abstract program level (software) that lends itself as an adequate analogy for the human mind. Programs are sets of instructions ready to act upon data; they may take diverse sequences of actions depending upon different conditions. As both instructions and data are stored in the same symbolic form, programs can also act upon themselves in a recursive manner, changing their own instructions if the task demands it.

When computers first appeared, they were largely considered to be mathematical devices, number-manipulating machines that would make calculations easier. But it soon became apparent that they were able to deal not only with numbers but also, more generally, with symbols. Thus, computers came to be conceptualized as symbol-manipulating systems. Computers can take symbolic input, recode it, store it, retrieve it, make decisions based on the recoded information, and produce symbolic output. These operations closely resemble mental processes and offer a precise and mechanistic way of thinking about them. Hence the computer became the best metaphor to lead the scientific research of the human mind.

There is general agreement that the critical time period giving a boost to cognitivism took place between 1954 and 1960. In 1956, a summer seminar took place at Dartmouth, where Allen Newell, Herbert Simon, and other major workers on artificial intelligence gathered to establish a research program for the new discipline interested in building programs able to generate intelligent behavior. In 1958, another summer seminar was organized by Newell and Simon at the RAND Corporation in Santa Monica in order to acquaint psychologists with computer simulation techniques and their application to the study of cognitive processes. In 1960, the Center for Cognitive Studies was created at Harvard, where Jerome Bruner and George Miller served as a point of reference for the new cognitive psychology.

3. COGNITIVE SCIENCE

The time from 1960 to approximately 1976 was a consolidation period for cognitive psychology. Processes such as attention, memory, language, and thought moved to the forefront of psychological research. In 1967, Ulric Neisser published a textbook, titled *Cognitive Psychology*, that served as a reference for the psychological research that was being done within the new cognitive framework. The heuristic value of the computer analogy was widely accepted, though it was

not understood the same way by all researchers. For some, the computer was just a metaphor that helped to conceptualize human mind as an information-processing system. For others, however, the computer was much more than a metaphor; they assumed that both the computer and the mind were examples of a type of system with specific properties. In 1976, Newell and Simon proposed the term physical symbol system as a name for this type of system and as a fundamental concept for a general and unifying approach to cognition that came to be known as cognitive science. That same year, the Sloan Foundation became interested in this approach and created an interdisciplinary research program to explore its possibilities. In 1977, the journal *Cognitive Science* was founded, and the Cognitive Science Society followed 2 years later. Nowadays cognitive science has been established as an academic discipline in several universities in different countries.

Cognitive science was an interdisciplinary enterprise in which different branches of scientific knowledge dedicated to the study of cognition merged. Originally, there were six disciplines involved, namely, psychology, computer science, linguistics, philosophy, anthropology, and neuroscience. Each had developed a particular way of looking at cognition, and some already had productive bilateral relations that had given rise to new fields of cognitive research, as was the case for psycholinguistics or computational linguistics. Nevertheless, there was the shared belief that progress within each discipline largely depended upon progress in the rest of them. Today, links between the six disciplines exist, though some are stronger than others, and new fields of research, as is the case for cognitive neuroscience, have emerged.

4. TENETS OF COGNITIVISM

4.1. Levels of Explanation

Cognitivism is not reductionist; it maintains that complex information-processing systems, such as the computer or the human mind, cannot be understood as a simple extrapolation of the properties of its elementary components. To understand a complex system, we have to take into consideration the fact that its behavior can be described at different levels of abstraction, each related to particular questions that may be asked about the system. If an explanation is conceptualized as an answer to a specific question, then we must be open to expect different explanations depending upon the description level associated to the question. The classical view of cognitive science distinguishes at least three different levels of analysis, named according to David Marr terminology.

4.1.1. The Computational Level

The computational level is the level of abstract problem analysis and refers to what the system does and why. At this level, explaining why people or computers do something requires pointing to their goals and objectives and to the strategies and means to carry them out. Take, as a simple example, the case of a calculator able to perform number addition. An explanation of what the machine does and why should be found in arithmetic theory, because the task to be performed by the system has to comply with the principles of this theory.

4.1.2. The Algorithmic Level

The algorithmic level is the program level in which a representation for coding the system's input and output should be specified, together with an algorithm providing the appropriate output for a given input. This level refers to how the system brings about its task. It is also called the symbol level, because the goals established at the computational level are encoded by symbolic expressions. In the case of the calculator from the example above, explaining how the system resolves an addition has to do with the representation used to encode numbers (Arabic, Roman, or binary, for instance) and the specific algorithm employed, for example, adding the units first and carrying to the tens if the sum exceeds 9.

4.1.3. The Level of Physical Implementation

The level of physical implementation is concerned with the particular technology used to physically realize the system. In our example, an explanation at this level would point to the working of electronic circuits, but electronic circuits are not the only way to implement an addition machine. Old cash registers, for example, were not electronic but mechanical; however, they could use the same algorithm as an electronic calculator to perform addition.

These three levels are frequently related and influence each other; for example, the choice of electronic circuits to build an addition machine may favor a binary over a decimal representation of numbers. However, it is convenient to note that, because the

three levels are not necessarily bound, some phenomena may find an explanation at only one or two of them. Above all, it should be noted that it is pointless to search for only one valid explanation of the behavior of complex information-processing systems. An explanation only makes sense in relation to a question referred to a particular level of analysis. Therefore, we need explanations at different levels in order to understand complex information-processing systems.

4.2. Systems Approach

Cognitivism views human cognition as an activity emerging from the interaction of a system of components. Though the search for components is considered a valuable aspect of the scientific activity, cognitivism emphasizes the significance of structure and functional architecture for our understanding of the mind.

From the beginning, cognitive psychology made wide use of flowchart models as a means of specifying the component processes involved in a particular situation. D. Broadbent's model was one of the first to specify the attention and memory components needed to explain human behavior in situations in which a person has to select one out of several different streams of information, as exemplified by a cocktail party, where a person has to follow one among many surrounding conversations. At first, models tended to be bound to specific tasks or situations, but soon interest for building general models became dominant. The model of human memory advanced by R. Atkinson and R. Shiffrin in 1968 is a good example of this trend. These authors were able to develop a memory model consisting of three different components, namely, sensory registers, short-term memory, and long-term memory. Extensive previous research had been carried out on each of these components, but Atkinson and Shiffrin managed to put them together within a general framework that was influential for years. Memory research also offers a good illustration of the analytic procedure that cognitive research employed to determine the components of a cognitive system. Five years after the model of Atkinson and Shiffrin was published, E. Tulving argued in favor of the partition of long-term memory into an episodic component and a semantic component. Later, in the 1980s, it was proposed that both episodic and semantic memory were part of a declarative or explicit component of memory having to do with conscious knowledge, but research also focused on an implicit, largely unconscious component of memory dealing with knowledge

related to cognitive, motor, and perceptual skills. Though there is no general agreement about the psychological reality of all these components and some authors consider the increasing number of them as a plain violation of the scientific principle of parsimony, a large group of psychologists views this tendency as a result of many years of analytic research and as a valuable way to make contact with brain research.

With the advent of cognitive science, the term cognitive architecture came to be used as a way to indicate that the structure of the cognitive system had some sort of primitive and permanent character. Search for this character divided the field of cognitive science into two conflicting views. On the one hand, the classical view considered the concept of the physical symbol as the building block of cognitive architecture and defined a cognitive system as a physical symbol system. On the other hand, a new connectionist view considered the concept of the neural network to be the basic unit of analysis for understanding the human cognitive system.

A physical symbol is a physical entity able to designate or refer to another entity that is the meaning of that symbol. For centuries this capacity to refer to something else was thought to belong only to the realm of ideas, and an idea was considered to be nonmaterial. As mentioned previously, the computer became a model for the human mind when scientists became aware that a computer was a symbol-manipulating machine rather than a machine dealing only with numbers. The physical symbol concept, by attributing to physical entities the capacity of having meaning, became a key concept in bridging the gap between brain and mind. In turn, a physical symbol system was defined as a system constituted by physical symbols. The adaptive control of thought (ACT) and the Soar systems, developed by John Anderson and Allen Newell, respectively, are two of the most influential current symbolic architectures.

The connectionist view represented an attempt to associate cognitive science with brain theory. Though the origin of the concept neural network goes back to Donald Hebb, and some interesting work on neural networks had already been done during the first years of cognitivism, the impulse that brought connectionism to the fore took place in the mid-1980s. Connectionists tried to build information-processing systems that were neurally inspired and called computation on such a system brain-style computation because they thought that the primitive concepts of a cognitive model should resemble the primitive units of the brain, neurons.

A neural network is composed of a set of elementary, neuron-like processing units connected to each other

in a specific way. The different strengths of the connections among units defines a pattern of connectivity that may change through experience according to a particular learning rule. An activation rule determines the way different inputs to a unit combine to determine the unit's state of activation, and an output function is in charge of mapping the unit state of activation into output. The only other component of a neural network is an environment within which the system can operate. Systems of this kind have been successfully used to explain performance on cognitive tasks as well as to model cognitive processes. For some time there were major disputes between defenders of symbolic and connectionist architectures. Nowadays, the arguments have abated, and many cognitive scientists make free use of both symbolic and neural network architectures depending upon the problem at hand. Neural networks seem best suited to model knowledge dependent on long, repetitive practice, as is usually the case in implicit knowledge. On the other hand, symbolic architectures appear preferable for modeling high-level cognitive processes such as reasoning or thought.

4.3. Representation

One of the main characteristics defining a particular system as cognitive is its capacity to reinstate a prior experience in the absence of a current external stimulus. We say that such a system owns the capacity to represent its environment.

Interest in mental representation as such was not a relevant feature of the first cognitive models. As discussed previously, they mainly dealt with mental structure and its components. Reference to mental representation was implicit in the use of terms such as encoding and recoding, but interest in the mental code itself had to wait until the early 1970s.

To understand what representation is we should differentiate the representing world from the represented world. The representing world has to somehow imitate the represented world, though not every possible aspect of the latter must be reflected in the former. For any representation, the following three features must be specified:

1. What aspects of the represented world are chosen to be encoded;
2. What elements of the representing world are doing the encoding;
3. What the correspondence is between the two worlds.

Research on mental images and memory brought to the forefront the very problem of mental representation. Was the representation responsible for mental images different from abstract semantic or syntactic representations? Were there analogical representations? These questions divided the field of cognitive psychology in two halves, one for and the other against analogical representations, a debate that still continues, although recent research in cognitive neuroscience seems to favor their existence. In the field of semantic memory, a plethora of representational formats for abstract knowledge were advanced within the framework of symbolic architectures. Prototypes, semantic features, semantic networks, schemata, and production systems were some of the main formats employed, but none attained general acceptance. Some investigators tried to build semantic representations based on a handful of primitive structures, but most models using symbolic architectures turned to local representations. In a local representation, there is a one-to-one correspondence between entities in the represented world and elements in the representing world.

Connectionist systems use distributed representations. In a distributed representation, each entity is represented by a pattern of activity distributed over many elements in the representing world, and each element represents many different entities. Distributed representations have proved to be very effective in dealing with aspects of cognition that were difficult to implement on symbolic architectures. Content-addressable and reconstructive memories, ability to generalize automatically to novel situations, and capacity to be modified according to a changing environment are genuine achievements of distributed architectures.

Cognitive science, symbolic cognitive science in particular, has been criticized for the way it has elaborated the notion of representation. Criticisms range from denying a particular characteristic of the representation, like that of being symbolic or abstract, to calling the entire notion into question. A common trait for these views is their concern about treating mind, body, and environment as coupled systems, dynamically interacting with each other; hence the term situated cognition is often used for this approach. It emphasizes that representations undergo continuous changes in tune with changes in the environment and encourages research on the changing patterns of behavior over time, using nonlinear differential equations to do the modeling. From this point of view, the best way to construct cognitive models is to build real agents. So far, the main contributions of this viewpoint

are to be found in the fields of perception and motor processes. To what extent this approach may be useful in understanding higher cognitive processes remains to be seen.

4.4. Process

Processes are operations acting on a representation either to generate it or to transform it in some way. They may be involved in translating external input into a particular representation, modifying the representation itself, or generating output. A process may be considered a function mapping an input to an output.

Interest in elementary mental operations goes back to the Dutch physiologist F.C. Donders in 1868. He assumed that the time spent by an agent performing a task could be divided into a series of stages, each of which corresponded to an elementary component process of the task. By using simple tasks, supposedly consisting of few elementary components, and subtracting the time spent in performing one task from the time corresponding to the next more complex one (with an extra elementary component), Donders thought to have isolated the time for simple cognitive processes such as detection or discrimination. Two assumptions relevant for cognitive psychology underlay Donders' experimental program: the existence of elementary mental operations and their serial organization as component processes of a complex task.

The information-processing approach has tended to rely on the assumption that a small number of elementary mental operations are sufficient to specify the complexity of human cognitive performance. However, there is little agreement about what the elementary operations are and about the ultimate level of description appropriate for an elementary operation. Connectionist models offer an alternative because they use only activation and inhibition as elementary processes acting on neuron-like elementary units, but many researchers prefer descriptions in more intuitive terms referring to information-processing stages.

A central issue for information-processing theories is the way in which mental processes are assembled in order to perform a task. Early cognitive models assumed a serial organization of processes; they were carried out one after another, each process taking as input the output of the previous one, and starting operation when the preceding process had finished. It soon became apparent that many cognitive processes in perception, memory, and other cognitive areas could not be understood on the basis of serial organization only; rather, they seemed to operate simultaneously, that is, in parallel. Parallel functioning of processes is typically associated with division of labor and modularity. A module, much like a subroutine in a computer program, is made up of a set of operations carrying out a particular assignment in the service of a larger task. It has been much debated whether the entire human mind or only some cognitive processes are modularly arranged. Nevertheless, there seems to be general agreement that modularity is a characteristic of low- rather than high-level cognitive processes. Connectionism has made parallel processing a distinctive feature of its approach. The term parallel distributed processing, often used as an alternative name for connectionism, highlights the two defining characteristics of this view, namely, distributed representation and parallel processing. Even so, many connectionist models group their simple, neuron-like units into layers, thus allowing for some serial organization of information processing stages.

An additional dimension introducing differences within cognitive theory has to do with the direction of the processing flow. Some theories and models emphasize bottom-up processing, in which the entire job is carried out by algorithms acting on input data to build increasingly complex representations. David Marr's influential theory of vision is a good example of a cognitive theory built from a bottom-up point of view. On the other hand, top-down processing emphasizes the feedback influence that complex cognitive component processes, such as goals, beliefs, and knowledge, exert on the functioning of simpler ones. A process is said to be data driven when its working is independent of feedback influences from goals, intentions, or other higher level processing stages. A process depending upon these influences is said to be conceptually driven. The most powerful cognitive models are usually interactive, meaning that they include both bottom-up and top-down processing.

Another frequently used classification of cognitive processes is the division between automatic and controlled processes. An automatic process usually refers to a sequence of component subprocesses that is triggered by a particular input. The sequence may include both bottom-up and top-down processing. An automatic process may be built-in as part of the physical implementation of a system, or its components may be assembled as a result of practice. Thus, to the extent that automaticity depends on practice, it should be considered a gradual rather than an all-or-none property for processes. On the other hand, when a new sequence of component

processes is required, top-down attentional influences are needed to do the assembling. In this case, the sequence is called a controlled process, meaning that it is under attentional control. Automaticity and dependence upon attentional control are complementary characteristics of mental processes; the more automatic a process is, the less attentional control it needs to act. Distinguishing between automatic and controlled processes has become very useful in psychological research. This distinction partially overlaps with that between explicit and implicit or between conscious and unconscious processes. Further research on this area will likely reveal fundamental characteristics of the human cognitive system.

4.5. Methods

Cognitivism brought to psychology new research methods and new ways of approaching more traditional ones. Reaction time came to be one of the main dependent variables used to understand cognitive processes. As mentioned before, Donders' subtractive method and more sophisticated later developments, like S. Sternberg's additive factors method, served as a guide in that reaction time was not studied by itself, but as an index for making inferences about nonobservable processes. Accuracy measures were also developed with the same purpose. In 1954, signal-detection theory was applied to perception by W. Tanner and J. Swets, providing indexes of perceptual functioning uncontaminated by the observer's decision biases. Subsequently, signal detection was also applied to memory and became a valuable tool for studying diagnostic systems.

Modeling cognitive processes was a research method characteristic of the cognitive approach. Mathematical as well as flow-chart models were widely used in the early days, but computer models, either symbolic or connectionist, were soon considered the high road to understanding cognitive processes. Computer modeling helped psychologists become aware of cognitive processing intricacy and forced them to be specific about the mechanisms involved in a particular task.

Recently, methods from the realm of neuroscience have been incorporated into cognitive research. These days, electrophysiological and brain-imaging techniques allow online recording of brain activity while a person is performing specific cognitive tasks. These technological advances provide a unique opportunity to relate brain activity to cognitive performance and to establish connections between cognitive constructs and brain structures and processes. An entirely new and promising field of research, known as cognitive neuroscience, has appeared, bringing new horizons for cognitive research.

5. COGNITIVIST INFLUENCE

The influence of cognitivism pervades every field of psychological research and spreads to the entire realm of social sciences. The term cognitive, considered vague and nonscientific during the behaviorist era, has now become respectable and scientific. Different approaches within cognitive science provide new ways of thinking about mental processes and intelligence. The study of human intelligence, once restricted to the statistical analysis of psychological tests, has now expanded to the working of mental machinery. The intelligence factors isolated by means of factor analysis are being studied in relation to cognitive constructs, thus bridging a long-lasting gap between differential and experimental psychology. As almost every article in this encyclopedia shows, cognitive factors are frequently mentioned as relevant causes in the explanation of different practical situations. We learn about cognitive ergonomics, cognitive behavior modification, cognitive neuropsychology, and other cognitive enterprises as truly influential approaches within applied psychology. Each one deserves to be treated separately.

See Also the Following Articles

Classical Conditioning ■ Cognition and Culture ■ Cognitive Skills: Training, Maintenance, and Daily Usage ■ Connectionism

Further Reading

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes, In K. W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation* (pp. 89–195). New York: Academic Press.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Clark, A. (1997). *Being there: Putting brain, body and world together again.* Cambridge, MA: MIT Press.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution.* New York: Basic Books.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (1998). *Cognitive neuroscience: The biology of the mind.* New York: W. W. Norton & Company.

Gruber, H. E., & Vonèche, J. J. (1977). *The essential Piaget.* New York: Basic Books.

Lachman, R., Lachman, J. L., & Butterfield, E. C. (1979). *Cognitive psychology and information processing: An introduction.* Hillsdale, NJ: Lawrence Erlbaum.

McClelland, J., Rumelhart, D., & Hinton, G. (Eds.) (1986). *Parallel distributed processing: Explorations in the micro-structure of cognition, Volume 1: Foundations.* Cambridge, MA: MIT Press.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.

Posner, M. I. (Ed.) (1989). *Foundations of cognitive science.* Cambridge, MA: MIT Press.

Posner, M. I., & Raichle, M. E. (1994). *Images of mind.* New York: Scientific American Library.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard University Press.

Collective Action

Bert Klandermans

Free University of Amsterdam, Amsterdam, The Netherlands

GLOSSARY

collective action frames Sets of collective beliefs that serve to create a frame of mind in which participation in collective action appears meaningful.

collective identity A place shared with other people.

collective incentives Incentives that relate to achievement of the movement's goals and the extent to which participation in a specific activity contributes to goal achievement.

erosion of support The process whereby people lose their sympathy for the movement in the course of a mobilization campaign.

free rider A person who refrains from participation, assuming that he or she will reap the benefits of the collective action anyway.

grievances Experiences of illegitimate inequality, feelings of relative deprivation, feelings of injustice, and/or moral indignation about some state of affairs.

identity Participation as an expression of identification with a group.

ideology Participation as a search for meaning and an expression of one's views.

instrumentality Participation as an attempt to influence the social and political environments.

mobilization The transformation of potential participants (sympathizers) into participants.

nonconversion of support The failure to convert sympathizers into participants because organizers failed to target them or were unable to motivate them.

personal identity The set of places that a person occupies in society.

selective incentives Incentives that relate to costs and benefits that differentiate between participants and nonparticipants.

On February 15, 2003, millions of people took to the streets in hundreds of cities in 60 countries throughout the world. In what was probably the largest collective action ever staged in human history, they protested against the imminent war in Iraq. Men and women of all ages expressed their aversion to the planned invasion. They were mobilized through the media, through organizations in which they were members, and by people who knew them. They came alone or with family and/or friends. They detested their government if it supported the war or applauded their government if it opposed the war. Obviously, these were not random samples of the population that walked the streets. The vast majority of the participants placed themselves at the left end of the political spectrum. They were predominantly white-collar workers or students who were highly educated. They were people who were interested in politics, and many had taken

part in demonstrations previously. The antiwar demonstrations were examples of an increasingly common form of political behavior, namely, collective action. What is collective action, why do people participate in such action, and how are these people mobilized? This article addresses these questions. Although the concept of collective action has a broader use, the article restricts itself to collective political action.

1. WHAT IS COLLECTIVE POLITICAL ACTION?

Collective political actions are challenges by people with common purposes and solidarity in interaction with elites and authorities. This definition includes two key elements that deserve some elaboration. First, collective political actions are challenges. They concern disruptive direct action against elites, authorities, other groups, or cultural codes. There is an obvious reason why this is the case. Collective political action is typically staged by people who lack access to politics. If they had access, there would be no need for the action. However, disruption forces authorities to pay attention to the claims brought forward. Second, collective political actions concern people with a common purpose and solidarity. Participants rally behind common claims. They want authorities to do something—to change a state of affairs or to undo changes. Such common claims are rooted in feelings of collective identity and solidarity. If these challenges gain sustainability, they constitute a social movement.

Collective political action has become much more frequent over recent decades. For example, in 1975, 22% of people in Britain signed at least one petition; in 1990, the figure was 75%. In 1974, 9% of people in Britain participated in at least one demonstration; in 1990, the figure was 25%. Between 1979 and 1993, the number of protest events in the French city of Marseille more than doubled, from 183 to 395 events per annum. During that same period, considerable increases in protest events in Germany were reported, a result that was confirmed in 1995 by Kriesi and colleagues' now classic study on new social movements in Europe. At the same time, diffusion to broader sectors of the population was witnessed. Across age groups and gender lines, from the left to the right, among workers and students, and in Western and non-Western societies alike, collective political action has become a common phenomenon.

2. WHY DO PEOPLE PARTICIPATE IN COLLECTIVE POLITICAL ACTION?

The social psychology of protest suggests three fundamental motives for participation in collective political action in that people may want to (a) change their circumstances, (b) act as members of their group, and/or (c) give meaning to their world and express their views and feelings. This article suggests that these three motives together account for most of the motivation to participate in collective political action in a society refers to these motives as instrumentality, identity, and ideology. Instrumentality refers to participation as an attempt to influence the social and political environment. Identity refers to participation as an expression of identification with a group. Ideology refers to participation as a search for meaning and an expression of one's views. This article does not suggest that these are mutually exclusive motives or are competing views on collective action participation, although some parties in the debates in the literature seem to have taken that position. However, this article concurs with those who argue that approaches that neglect any of these three motives are fundamentally flawed. To be sure, individual participants may participate due to a single motive, but all three motives are needed to understand why people take part in collective political action. Of course, that raises the question of how the three motives interact in fostering participation. The article elaborates on this in the following discussion of the motivation to participate.

2.1. The Motivation to Participate in Collective Action

2.1.1. Instrumentality

Instrumentality begins with dissatisfaction, be it the experience of illegitimate inequality, feelings of relative deprivation, feelings of injustice, moral indignation about some state of affairs, or a so-called suddenly imposed grievance, that is, an event imposed on people (e.g., highway construction in the neighborhood). Social psychological grievance theories, such as relative deprivation theory and social justice theory, have tried to specify how and why grievances develop. Despite the fact that grievances are at the roots of collective political action, they have not featured prominently in social movement literature since the early 1970s. Resource mobilization theory and political process theory, the

two approaches that have dominated the field since those days, have always taken as their point of departure that grievances are ubiquitous and that the key question in action participation research is not so much why people are aggrieved as it is why aggrieved people participate. The instrumentality paradigm holds that participants are people who believe that they can change their social and political environments at affordable costs.

The perceived costs and benefits of collective action can be distinguished in selective and collective incentives. Selective incentives relate to costs and benefits that differentiate between participants and nonparticipants. People are spending time and/or money only if they participate, people run the risk of being beaten up by the police only if they participate, people's friends will blame them for not participating only if they stay home, and so on. Sometimes, movement organizations try to make participation more attractive by providing selective benefits (e.g., a popular music group, a train ticket to the city where the demonstration is held, a T-shirt). Authorities or opponents, for their part, can try to make participation less attractive by imposing costs on participants. Collective incentives are related to achievement of the movement's goals and the extent to which participation in a specific activity contributes to goal achievement. Obviously, it is not enough for a goal to be important to a person. What is also needed is some likelihood of success. The problem with collective action is that it is difficult to know to what extent an activity will have any influence on authorities. In any event, chances are low that an activity will have any impact if only a few people participate. Therefore, the likelihood of success depends on the expected behavior of others. If too few people participate, it is unlikely that the activity will make any difference. As a consequence, expectations about the behavior of others play an important role in the decision to participate. If someone expects that only a few people will participate, that person's motivation to participate will be low. In a way, the expectation about other people's behavior functions as a self-fulfilling prophecy. If people believe that few people will participate, they will not be motivated to participate and, thus, will help make their own expectation come true.

Instrumentality presupposes an effective movement that is able to enforce some wanted changes or at least to mobilize substantial support. Making an objective assessment of a movement's impact is not an easy task, but movement organizations will try to convey the image of an effective political force. They can do so by pointing to the impact they have had in the past or to the powerful allies they have. Of course, they may lack all of these, but they might be able to show other signs of movement strength. A movement may command a large constituency as witnessed by turnouts at demonstrations, by membership figures, or by large donations. It may be composed of strong organizations with strong charismatic leaders who have gained respect. Instrumentality also implies the provision of selective incentives. Movements may vary considerably in the selective incentives for participation they provide. Obviously, this is also a matter of the resources that a movement commands. Surprisingly little comparative information is available on the resources that movements have at their disposal. In a similar vein, systematic documentation is lacking on the ways in which the larger political system and the alliances and opponents of movement organizations influence movement participation. In 1978, Tilly coined the now classic terms "repression" and "facilitation" to distinguish between political systems that increase and decrease the costs of participation. Indeed, repressive political environments may increase the costs of participation considerably. People may lose friends, they may risk their jobs or otherwise jeopardize their sources of income, they may be jailed, and they may even lose their lives.

Viewed from an instrumental perspective, a solution to the dilemma of collective action must be found. It was through the work of Olson that students of movement participation became aware of the problem. In 1968, Olson published his *The Logic of Collective Action*. The core of the book was the argument that rational actors will not contribute to the production of a collective good unless selective incentives persuade them to do so. Collective goods are characterized by "jointness of supply." That is, if they are made available to one person, they become available to everybody irrespective of whether people have contributed to the production of the collective good (e.g., a law against discrimination, measures against pollution). Therefore, according to Olson, a rational actor will choose to take a free ride unless selective incentives (i.e., costs or benefits that are made contingent on participation in the production of the collective good) prevent him or her from doing so. Olson's argument was soon applied to social movement participation. It helped to explain why so often people do not participate in social movements despite their interest in the achievement of the movements' goals. Movement scholars argued that movement goals typically are collective goods. If a goal

is achieved, people will enjoy the benefits irrespective of whether they have participated in the effort. In view of a goal for which achievement is uncertain but for which benefits—if they materialize—can be reaped anyway, rational actors will take a free ride.

However, social movement scholars quickly discovered that reality is more complex than Olson's simple model suggests. The problem with Olson's logic of collective action is that it provides an explanation for why people do not participate but fares poorly in explaining why people do participate. Moreover, Olson's solution that people participate for selective incentives is fundamentally flawed because it does not give a satisfactory answer to the question of where the resources needed to provide selective incentives come from. If these resources must be collected from individual citizens, the same collective action dilemma arises again. This is not to say that selective incentives are irrelevant; however, in the final instance, they cannot solve the collective action dilemma. In other words, if collective and selective incentives do not provide a sufficient explanation of movement participation, what else might make the difference? A recurring criticism was that Olson's model assumes that individuals make their decisions in isolation, that is, as if there are no other people with whom they consult, with whom they feel solidarity, and by whom they are kept to their promises. This pointed to the significance of collective identity as a factor in movement participation.

2.1.2. Identity

It soon became clear that instrumentality, or movement participation to achieve some external goal (social or political change), was not the only reason to participate. After all, many external goals are reached only in the long run if at all. Similarly, when it comes to material benefits, costs frequently outweigh benefits. Apparently, there is more involved in being a movement participant than perceived costs and benefits. Indeed, one of those motives relates to belonging to a valued group.

Identity can be described succinctly as a place in society. People occupy many different places in society. They are students, housewives, soccer players, politicians, farmers, and so on. Some of those places are exclusive, that is, occupied by only a small number of people (e.g., the members of a soccer team). Others are inclusive and encompass large numbers of people (e.g., Europeans). Some places are mutually exclusive (e.g., male/female, employed/unemployed), some are nested (e.g., French/Dutch/German vs European), and some

are crosscutting (e.g., female and student). All of these different roles and positions that a person occupies form his or her personal identity. At the same time, every place a person occupies is shared with other people. The author of this article is not the only professor of social psychology, nor is he the only Dutch or the only European. He shares these identities with other people, a fact that turns them into collective identities. Thus, a collective identity is a place that is shared with other people. This implies that personal identity is always collective identity at the same time. Personal identity is general, referring to a variety of places in society, whereas collective identity is always specific, referring to a specific place.

Most of the time, collective identities remain latent. Self-categorization theory hypothesizes that an individual may act as a unique person (i.e., display his or her personal identity) or as a member of a specific group (i.e., display one of the many collective identities he or she has depending on contextual circumstances). Contextual factors may bring personal or collective identity to the fore. Obviously, this is often not a matter of free choice. Circumstances may force a collective identity into awareness whether people like it or not. The Yugoslavian and South African histories have illustrated this dramatically, and there are other equally or even more dramatic examples throughout human history. But collective identities can also become significant in less extreme circumstances. For example, consider the possible effect of an announcement that a waste incinerator is planned next to a certain neighborhood. Chances are that within very little time, the collective identity of the people living in that neighborhood will become salient.

Self-categorization theory proposes that the more people identify with a category, the more they are prepared to employ a social category in their self-definition. Identification with a group makes people more prepared to act as a member of that group. This assertion refers, of course, to identity strength. Social identity literature tends to neglect that real-world identities vary in strength. But identifying more or less strongly with a group may make a real difference, especially in political contexts. Moreover, strong identities may be less affected by context. Following this reasoning, we may expect that strong identities make it more likely that people act on behalf of their group. The basic hypothesis regarding collective identity and movement participation is fairly straightforward. A strong identification with a group increases the likelihood of participation in collective political action on

behalf of that group. The available empirical evidence supports this assumption overwhelmingly.

Movements offer the opportunity to act on behalf of one's group. This is most attractive if people identify strongly with their group. The more farmers identify with other farmers, the more appealing it is to take part in farmers' protests. The more women identify with other women, the more inviting it is to participate in the women's movement. The more gay people identify with other gay people, the more they are attracted by the possibility of taking part in the gay movement. In addition to the opportunity to act on behalf of the group, collective political action participation offers the opportunity to identify with the movement's cause, the people in the movement, the movement organization, the group in which one is participating, and/or the leader of the movement. Not all of these sources of identification are always equally appealing. Movement leaders can be more or less charismatic, and the people in the movement or in someone's group can be more or less attractive. Moreover, movements and movement organizations may be, and in fact often are, controversial. Hence, becoming a participant in a movement organization does not mean taking a respected identity on oneself—quite the contrary. Within the movement's framework, of course, this is completely different. There, the militant does have the status that society is denying him or her. And for the militant, of course, in-group/out-group dynamics may turn the movement organization or group into a far more attractive group than any other group "out there" that is opposing the movement. Indeed, it is not uncommon for militants to refer to their movement organization as a second family, that is, as a substitute for the social and associative life that society is no longer offering them. Movement organizations not only supply sources of identification but also offer all kinds of opportunities to enjoy and celebrate the collective identity—marches, rituals, songs, meetings, signs, symbols, common codes, and so on.

2.1.3. Ideology

The third motive—wanting to express one's views—refers at the same time to a long-standing theme in the social movement literature and to a recent development. In classic studies of social movements, the distinction was made between instrumental and expressive movements or protests. During those days, instrumental movements were seen as movements that aimed at some external goal (e.g., implementation of citizenship rights). Expressive movements, on the other hand, were movements that had no external goals and for which participation was a goal in itself (e.g., expression of anger in response to experienced injustice). Movement scholars felt increasingly uncomfortable with the distinction because it was thought that most movements had both instrumental and expressive aspects and that the emphasis on the two could change over time. Therefore, the distinction lost its use. Recently, however, the idea that people might participate in movements to express their views has received renewed attention. Attention arose this time from movement scholars who were unhappy with the overly structural approach of resource mobilization and political process theory. These scholars began to put an emphasis on the creative, cultural, and emotional aspects of social movements such as music, symbols, rituals, narratives, and moral indignation. People are angry, develop feelings of moral indignation about some state of affairs or some government decision, and wish to make these feelings known. They participate in social movements not only to enforce political change but also to gain dignity in their lives through struggle and moral expression.

Ideology has a significant affective component. Acting on one's ideology is deemed to be one of the fundamental motives of action participation and is necessarily charged with emotion. Appraisal and action are socially constructed. That is, they are formed in interpersonal interaction, especially in the case of politically relevant emotions. Cultural and historical factors play an important role in the interpretation of the state of affairs by which politically relevant emotions are generated. Obviously, appraisal can be manipulated. Activists work hard to create moral outrage and anger and to provide a target against which these can be vented. They must weave together a moral, cognitive, and emotional package of attitudes. But emotions also play an important role in the ongoing activities of movements. The literature distinguishes two kinds of collective emotions—reciprocal emotions and shared emotions—that reinforce each other. Each measure of shared outrage against an injustice reinforces the reciprocal emotion of fondness for others precisely because they feel the same way. Conversely, mutual affection is a context in which newly shared emotions are easily created. Anger and indignation are emotions that are related to a specific appraisal of the situation. At the same time, people might be puzzled by some aspects of reality and may try to understand what is going on. They may look for others with similar experiences, and a social movement may provide an environment in which to exchange experiences, tell their stories, and express their ideologies.

Social movements play a significant role in the diffusion of ideas and values. In 1998, Rochon made the important distinction between "critical communities" (where new ideas and values are developed) and "social movements" (which are interested in winning social and political acceptance for those ideas and values). "In the hands of movement leaders, the ideas of critical communities become ideological frames," stated Rochon, who continues to argue that social movements are not simply extensions of critical communities. After all, not all ideas developed in critical communities are equally suited to motivate collective action. Social movement organizations, then, are carriers of meaning. Through processes such as consensus mobilization and framing, they seek to propagate their definitions of the situations to the public at large. A study of flyers produced by the various groups and organizations involved in the protests against the International Monetary Fund and the World Bank in Berlin, Germany, is an excellent example in this respect. The study showed how links were constructed between the ideological frame of the organizers of the demonstration and those of the participating organizations so as to create a shared definition of the situation. In the social movement literature, such definitions of situations have been labeled "collective action frames."

Social movements do not invent ideas from scratch; rather, they borrow from the history of ideas. They build on an ideological heritage as they relate their claims to broader themes and values in society. In so doing, they relate to societal debates that have histories of their own, and those histories are usually much longer than those of the movements themselves. In 1992, Gamson referred to the "themes" and "counterthemes" that, in his view, exist in every society. One such pair of a theme and countertheme that Gamson mentioned is that of "self-reliance" versus "mutuality," that is, the belief that individuals must take care of themselves versus the belief that society is responsible for its less fortunate members. Klandermans has demonstrated how, in The Netherlands, these two beliefs became the icons that galvanized debate and spurred protest over disability payments. Self-reliance became the theme of those who favored restrictions in disability payment, whereas mutuality was the theme of those who defended the existing system.

2.2. The Transformation of Potentiality into Action

People who are potentially prepared to take part in collective action do not automatically participate. Potential must be transformed into action. This is in fact what most mobilization campaigns are about—targeting potential participants and turning them into actual participants. Obviously, a campaign will never be 100% successful in that regard. In the course of a campaign, people may lose their sympathy for the movement. Klandermans has labeled this "erosion," a phenomenon that is more likely to occur if countermovements and/or authorities are mounting countercampaigns. On the other hand, organizers might fail to convert sympathizers into participants because they fail to target or motivate them. Klandermans has called this "nonconversion." The rest of this section focuses on targeting.

Targeting sympathizers involves finding answers to two strategic questions: "Who are the sympathizers?" and "How can they be reached?" Every attempt to persuade must find answers to these two questions. Indeed, sophisticated movement organizations such as Greenpeace have learned to employ many of the strategies that have been developed in advertising. Social networks are of crucial importance in this regard. Movement organizations have two options here in that they can (a) try to co-opt existing networks or (b) build new networks. Both strategies are mobilization efforts in themselves. Co-optation is the easier strategy of the two because it builds on existing commitments to organizations and networks that are part of the movement organization's alliance system. However, there are risks. The co-opted organizations may use the campaigns for their own ends, or their leadership may—for whatever reason—decide not to collaborate. The latter makes it more difficult for the rank-and-file to cooperate. Yet co-opting existing networks, such as churches, unions, political parties, and youth organizations, is frequently applied, if only because doing so implies an answer to both strategic questions. On the one hand, it works from the assumption that most members of the movement organization sympathize with the movement. On the other hand, it assumes that these sympathizers can be reached through the organization's networks.

Building new networks implies recruiting people who are willing to spend considerable amounts of time for a prolonged period as movement activists. Therefore, it will require more effort on the part of the organization than will co-opting existing networks. But once the new networks are established, they will be more reliable. The recruitment of movement activists is a process that is determined, on the one hand, by factors that influence who is being asked and, on the

other, by factors that influence who agrees to serve as activists when asked. As for the first type of factors, a crucial determinant is someone's position in networks linked to the movement organization or, more specifically, to the movement organizer who is undertaking the recruitment effort. Movement organizers tend to recruit first among the people they know, and this often suffices. After all, they do not need many activists. They need a few to maintain the networks, and once they have those few, the return on having additional activists diminishes rapidly. Indeed, it has been argued that long-term activism is one form of activism that must cope with free-rider behavior. The latter point is important when deciding whether to serve as an activist or not. Klandermans has maintained that the people who are asked to serve as activists understand perfectly well that they are giving a free ride to most of the sympathizers to the movement, but they are prepared to do so because they care. Only people who really care a lot are prepared to sacrifice for the others. They make the effort because they believe that "If [they] don't do it, nobody else will." Klandermans has asserted that only people who strongly support the movement's ideology, and who feel responsible for maintaining and proselytising the movement and its ideology, are prepared to make the effort of serving as movement activists.

See Also the Following Articles

Ideological Orientation and Values ■ Organizational Participation

Further Reading

Gamson, W. A. (1992). *Talking politics*. Cambridge, UK: University of Cambridge Press.

Klandermans, B. (1997). *The social psychology of protest*. Oxford, UK: Blackwell.

Klandermans, B., & Staggenborg, S. (Eds.). (2002). *Methods of social movement research*. Minneapolis: University of Minnesota Press.

Kriesi, H., Koopmans, R., Duyvendak, J. W., & Guigni, M. (1995). *New social movements in Western Europe: A comparative analysis*. Minneapolis: University of Minnesota Press.

McAdam, D., Tarrow, S., & Tilly, C. (2001). *Dynamics of contention*. Cambridge, UK: Cambridge University Press.

Olson, M. (1968). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Rochon, T. R. (1998). *Culture moves: Ideas, activism, and changing values*. Princeton, NJ: Princeton University Press.

Tarrow, S. (1998). *Power in movement: Social movements, collective action, and mass politics in the modern state* (2nd ed.). Cambridge, UK: Cambridge University Press.

Tilly, C. (1978). *From mobilization to revolution*. Reading, MA: Addison–Wesley.

Color Blindness

Julio A. Lillo and Humberto Moreira

Complutense University, Madrid, Spain

GLOSSARY

anomalous trichromatism Group of color blindness pathologies caused by disturbances in the functionality of a certain type of cone; there are three types of anomalous trichromatism: protanomaly (when the problem is related to the cones that are more sensitive to long wavelengths), deuteranomaly (when the alteration is related to the cones that are more sensitive to medium wavelengths), and tritanomaly (when the alteration is related to the cones that are more sensitive to short wavelengths).

color blindness Term used to name a set of color perception anomalies; the term is not totally accurate because most color-blind people do not have "black and white" vision but rather experience a more or less reduced gamut of chromatic sensations.

daltonisms (red–green pathologies) Term used to name the group of color blindness associated with the X chromosome; it includes two varieties of protan alterations (protanopia and protanomaly) and two varieties of deutan alterations (deuteranopia and deuteranomaly).

dichromatisms Group of color blindness pathologies caused by the absence of functionality in a certain cone type (usually, the cones do not have the proper photopigment); protanopia is the dichromatism related to problems in the cones that are more sensitive to long wavelengths, deuteranopia is the dichromatism related to the cones that are more sensitive to medium wavelengths, and tritanopia is the dichromatism related to the cones that are more sensitive to short wavelengths.

metamers Stimuli that, being physically different, produce an identical chromatic experience.

monochromatisms Group of color blindness pathologies; monochromats have functionality in only one specific type of photoreceptor, and they lack color perception.

opsin Protein present in the receptor photopigments.

scotopic vision The customary vision in normal-sighted people at night; it is mediated by the retina rods and does not allow color perception (and, consequently, is a monochromatic type of vision).

The scientific definition of color blindness is explained, emphasizing that this denomination is not only applied to people without color experiences (monochromats) but also to all with no-common color experiences (dichromats and anomalous trichromats). Clinical nomenclature in color-blind diagnosis is described, specifying the causal origin of each color-blind type. The different action types available to applied psychologists to help color-blind people are described (diagnosis, perceptual problems detection, use of specific information, creation-selection of accurate materials, etc.).

1. INTRODUCTION

The existence of metamer stimuli (i.e., physically different but perceptually identical) has psychological implications that are essential to understand the functional logic of color perception in both normal-sighted and color-blind people. Figure 1 shows the spectracolorimetric measurement of two metamers. The upper portion (Fig. 1A) is a physical description of the light reflected by a white shirt. The middle portion (Fig. 1B) corresponds to the light from this object as reproduced on a conventional television screen. It is easy to observe considerable differences between these two physical stimuli that, nonetheless, produce the same perceptual experience ("white").

The existence of metamers casts serious doubts on what could be called the "commonsense theory of color perception and color blindness." This theory can be summarized in the following three ideas:

1. Color is a physical property of light (e.g., the light reflected by blood is red, the light reflected by a lawn is green).
2. The function of our visual system consists exclusively of detecting such properties from the stimulation that reaches the retina.
3. Color blindness would derive from the visual system's limitations to detect the presence of some color (e.g., difficulty in detecting the presence of red).

From the viewpoint of the commonsense theory, the existence of metamers implies that all humans are color-blind because they experience the same color in response to physically different stimuli. It must be emphasized that such stimuli generate different color experiences when presented to other animal species (e.g., the stimuli from Figs. 1A and 1B are not metamers for bees and pigeons). Consequently, it can be concluded that the essential difference between normal-sighted and color-blind people is quantitative and not qualitative: All of them have metamers, but color-blind people have more or them. Although this conclusion is correct, a different conclusion can be inferred regarding the main premise of the commonsense theory because color is not a physical property of light. Instead, it is a property of the perceptual response generated in reaction to light, as was explicitly stated by Isaac Newton in his book *Optics:* "The rays, to speak properly, are not coloured. In them there is nothing else than a certain power and disposition to stir up a sensation of this or that colour."

Following Newton's ideas, the next section reveals the nature of the "certain power and disposition" present in physical stimuli and, even more important, how our visual system produces "a sensation of this or that colour" in response to them. Subsequently, it will be easy to conclude that the most important difference between normal-sighted and color-blind people is that the latter have a reduced capacity to detect energetic imbalances in light and, consequently, to produce color experiences.
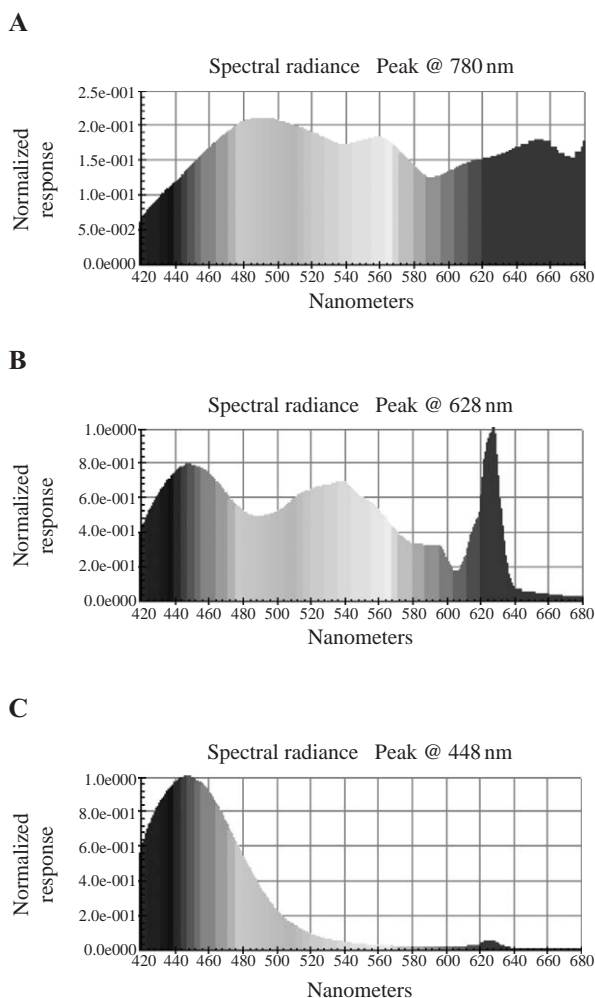
A

B

C

FIGURE 1 Spectracolorimetric measurements of three stimuli. Horizontal axis indicates the wavelength in nanometers (1 nm = 10^{-9} meters). Vertical axis indicates the relative quantity of energy. (A) White shirt. (B) White shirt on a television screen. (C) Primary blue screen.

2. USING THE ANOMALOSCOPE TO DIAGNOSE: PHOTORECEPTORS AND TYPES OF COLOR BLINDNESS

Color perception is the result of a complex visual processing sequence, initiated in the retina photoreceptors, that involves a large number of brain centers. Consequently, any anomaly in this sequence can generate some type of color blindness. However, most of the problems related to color perception are caused by some functional anomaly in a photoreceptor type. Consequently, this analysis centers on these anomalies.

There are two types of photoreceptors in the human retina: cones and rods. Cones are especially relevant to understanding color perception and can be divided into three types (continuous curves in Fig. 2), depending on the relative position of their spectral sensitivity curves. In the standardized scientific nomenclature, these types are called "L" (long), "M" (medium), and "S" (short), depending on the portion of the spectrum that produces larger responses for each type of cone. Taking into account the relative curve position and counting from right to left, the clinical nomenclature calls the three cone types "proto," "deutera," and "trita," respectively (corresponding to the Greek words for "first," "second," and "third"). Recent genetic evidence indicates that, at least for the L and M cones, the spectral response curves

present minor interindividual differences among normal-sighted people (discontinuous curves in Fig. 2).

The analysis of cone responses is useful to understand (a) the origin of metamers, (b) their more widespread presence among color-blind people and the various types of color blindness, and (c) how they could be diagnosed using an anomaloscope. Each of these aspects is considered in turn.

2.1. The Origin of Metamers

Despite their physical differences, the two metamers in Figs. 1A and 1B share an essential characteristic: They both present important energy accumulations in the three spectrum portions (long, medium, and short). That is why both stimuli generate high responses in the three cone types. Cone response similarity is the cause of the achromatic sensation generated ("white"). On the other hand, the stimulus represented in Fig. 1C accumulates energy in only one spectrum portion (the short one) and, consequently, generates an imbalance in the relative responses of the three cone types (the S cones have a larger response). This imbalance causes a chromatic sensation ("blue"). To summarize, first, the energetically unbalanced stimuli (more energy in any spectrum portion) tend to generate imbalances in the relative responses of the three cone types and, consequently, the generation of hue
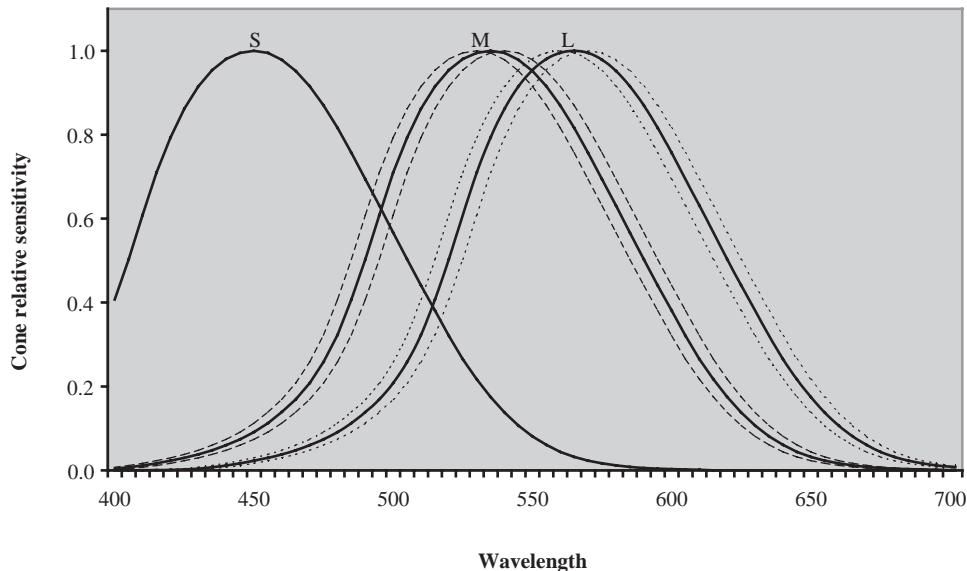

FIGURE 2 Cone spectral sensitivity curves. Horizontal axis indicates the wavelength in nanometers. Vertical axis indicates the relative sensitivities of the cones. S = cones most sensitive to short wavelengths; M = cones most sensitive to medium wavelengths; L = cones most sensitive to long wavelengths.

experiences (e.g., blue, green, purple). Second, the specific hue experienced depends on the type of imbalance detected by the cones. Third, all physical stimuli that generate the same pattern of response in the cones are metamers. For example, all stimuli that produce the same predominance of the S cones over the other cones will generate the same blue perception, and all stimuli that produce high and similar responses in the three cone types will generate the same white perception.

People with only one photoreceptor type lack the ability to detect energetic imbalances and are called "monochromats." Their vision is very similar to that experienced by normal-sighted people when watching a black and white film. The psychophysical characteristic of these individuals is that, by performing the necessary intensity adjustments, they can experience any pair of wavelengths as metamers. Equation (1) synthesizes the cause of this fact. As can be observed, the response (R) in the type of cone available depends on the relative wavelength visual efficacy (EF_λ) and physical intensity (I_λ) of the stimulus:

$$R = EF_\lambda \cdot I_\lambda. \tag{1}$$

2.2. Metamers and Types of Color-Blind People

The term "scotopic vision" refers to the kind of vision mediated by rods. When the illumination level is reduced and the cones become inoperative, vision depends on the rods and all humans become monochromats because they see with only one type of photoreceptor. Consequently, all of the variations have a quantitative characteristic and can be described using Eq. (1). Monochromatic vision is the only type of vision available to people who lack cones (rod monochromats). In addition, these individuals have severe acuity problems (similar to normal-sighted people at night, they cannot detect subtle details) and must use sunglasses permanently to reduce the amount of light reaching their eyes. Also considered monochromats are individuals with only one cone type (the S cone), although they do not have photophobia or acuity problems.

Strictly speaking, only monochromats should be called "color-blind" people because they are the only ones whose visual world is similar to that of normal-sighted people when watching a black and white film. However, monochromats comprise only a small portion of the people who are usually called color-blind.

People lacking a certain type of cone are dichromats. In the clinical taxonomy, this condition is indicated by using the suffix "anope" after the cone denomination. Thus, people lacking L cones (protocones) are protanopes, people lacking M cones are deuteranopes, and people lacking S cones are tritanopes. To reach a better understanding of dichromats' color limitations, a brief description of the most prestigious diagnostic test, the Nagel analytic anomaloscope, is presented.

2.3. Diagnosing Dichromatism and Anomalous Trichromatism with an Anomaloscope

The Nagel anomaloscope uses three monochromatic stimuli of 546, 589, and 670 nm (vertical lines in Fig. 3). These values are out of the operative range of the S cones (not represented in Fig. 3). Consequently, the next analysis can focus on the other two cone types: the M cones (thick left curve in Fig. 3) and the L cone (thick right curve in Fig. 3).

The anomaloscope shows a circle divided into two halves. One half is used to present the reference stimulus, a monochromatic stimulus with a wavelength of 589 nm (represented by the central vertical line in Fig. 3). To make the explanation simpler, assume that presentation of the reference stimulus generates similar activation levels in both cone types (L and M) and that, because of this, a yellow experience is generated. The second half of the circle presents a mixture of 546- and 670-nm light (represented, respectively, by the left and right vertical lines in Fig. 3). For a normal-sighted person, a 546-nm monochromatic stimulus generates more response in the M cones than in the L cones and, consequently, the perception of green. On the other hand, the 670-nm monochromatic light is experienced as reddish because it generates a larger response in the L cones than in the M cones.

The use of a Nagel anomaloscope requires adjusting the quantities of both mixture components (the greenish and the reddish light) to match the reference stimulus. For a normal-sighted person, this means that the mixture should similarly activate the M and L cones and, therefore, produce the same yellow experience as does the reference stimulus. These adjustments are performed very precisely because every variation in the relative quantities of the mixture components will produce a greenish or reddish hue and, consequently, an easily noticeable difference in relation to the reference color.

When carrying out Nagel anomaloscope adjustments, the more common dichromat types (protanopes and
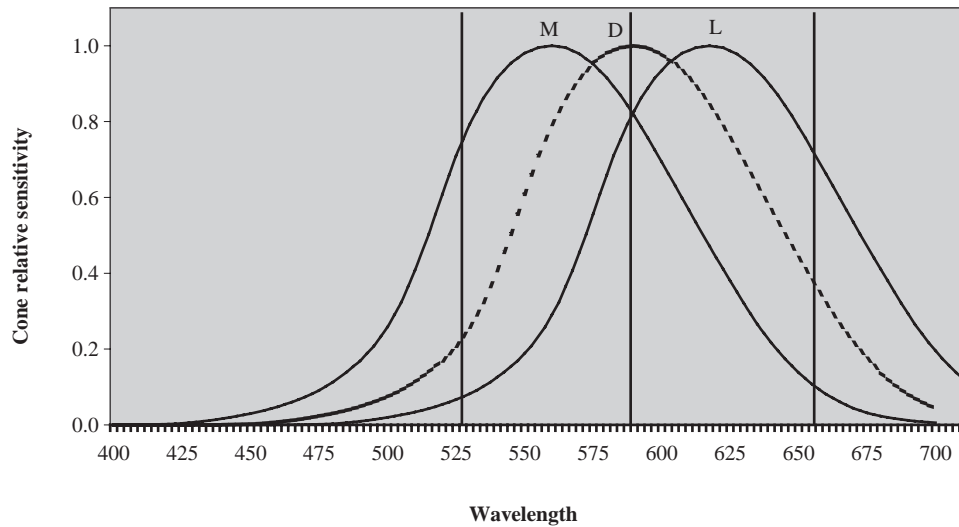
FIGURE 3 Anomaloscope stimuli and cone spectral curves. Horizontal axis indicates the wavelength in nanometers. Vertical axis indicates the relative sensitivities of the cones. M = cones most sensitive to medium wavelengths; L = cones most sensitive to long wavelengths; D = curve that substitutes one of the two previous ones in anomalous trichromats.

deuteranopes) are not as consistent as normal-sighted people. This is because the three stimuli used act on the same cone type, that is, the only one they have (remember that S cones are irrelevant to this analysis). Consequently, the situation is very similar to the one described by Eq. (1); the response (R) generated by the reference stimulus can be metamerized by any of the two mixture stimuli presented individually, or by any combination of both mixture stimuli (given the accurate quantities).

Anomalous trichromats are the next group of color-blind people to be considered. This group consists of all people who, like normal-sighted people, have three cone types in their retinas (this is why they are trichromats) but whose spectral curve responses are somewhat different from those represented in Fig. 2 (this is why they are anomalous). The most important consequence of this is that anomalous trichromats perceive as metamers certain stimuli perceived as different by normal-sighted people.

Nowadays, it is commonly assumed that the anomalous trichromat's limitations in differentiating colors, and the peculiar way in which these are perceived, derive from an exaggerated overlapping of the spectral response curves of two cone types. To understand why, the D curve in Fig. 3 must be used to substitute one of the two continuous curves (the M or L curve). With the new pair of curves, the relative differences in the response of both cone types to any stimulus are reduced, as is the ability to discriminate colors—the more overlapping there is, the less ability

to discriminate. At the limit, a full overlapping of curves would make both cone types functionally equivalent, and these people would behave like dichromats.

3. CAUSES OF COLOR BLINDNESS

Most color blindness is genetic in origin. In this group, the alterations associated with problems in the L cones (protanopia and protanomalies) or in the M cones (deuteranopia and deuteranomalies) are collectively called "daltonisms" or "red–green problems." These problems affect many more men (5–10%) than women (<0.9%) because they are associated with genetic alterations in recessively transmitted X chromosomes. Consequently, men will express red–green problems when there are anomalies in the only X chromosome they have. On the other hand, a women must have problems in both X chromosomes to be daltonic. With only one affected chromosome, they will show a normal use of colors but can transmit the red–green disturbance to their offspring.

The most important gene in relation to the opsin characteristic of S cones is not located in sexual chromosomes; instead, it is located in the seventh pair of chromosomes. Consequently, problems associated with these cones have a similar incidence in men and women. These types of color blindness are transmitted dominantly, but they affect a very small number of people (<0.05%).

TABLE I
Comparison of Acquired and Congenital Color Vision Deficiencies

Acquired color vision defect	Congenital color vision defect
Onset after birth	Onset at birth
Monocular differences in the type and severity of the defect occur frequently	Both eyes are equally affected
Color alterations are frequently associated with other visual problems such as low acuity and reductions in the useful visual field	The visual problems are specific to color perception; there are no problems with acuity (except in rod monochromatism) or visual field
The type and severity of the deficiency fluctuate	The type and severity of the defect are the same throughout life
The type of defect might not be easy to classify; combined or non-specific defects occur frequently	The type of defect can be classified precisely
Predominantly either protan or deutan	Predominantly tritan
Higher incidence in males	Same incidence in both sexes

Although the chromosomes associated with rod monochromatism have not yet been identified, this type of problem is known to be transmitted recessively. Consequently, the affected gene must be located in both chromosomes to produce problem expression.

Defective color vision can be acquired by any kind of agent with the capacity to disturb the normal functioning of the visual system—the result of ocular pathology, intracranial injury, the use of therapeutic drugs, aging, and so on. Frequently, color disturbances are initial symptoms of pathologies and can be used for early detection, indicating when therapeutic measures should begin or when they should be discontinued. Table I reviews the most important differences between acquired and congenital color deficiencies.

4. COLOR BLINDNESS AND PSYCHOLOGICAL INTERVENTION

4.1. Problems and Solutions: A Case of Protanopia

Mr. X works as an editor for a famous Spanish scientific dissemination journal. This activity is not severely handicapped by color blindness, according to various studies. Mr. X was diagnosed as a protanope some years ago. Although his perceptual limitations made it difficult for him to perform tasks in the same way as other people do, he had no problems after carrying out some modifications in his work environment. His past difficulty in correcting page galley proofs and the way

in which he "solved" this problem are an example. This protanope used to receive printed pages on which the errors to be corrected were indicated by red pencil marks. For a normal-sighted person, locating red marks is a very easy task because they "pop up" over the black letter background. However, for Mr. X, dark red and black colors are very similar (nearly metamers), so red marks did not "pop up" from his perspective. Consequently, he missed many error marks. How did he prevent this situation? Very simply: by using a color that would stand out over the background of black letters. In his case, blue ink was a very appropriate solution.

The kind of difficulty described in the previous paragraph has been experienced for centuries. To avoid them, people usually adopt one of two alternatives. Either color-blind people are denied access to some tasks and/or activities, or some task modifications are carried out to make them accessible to color-blind people. In any case, neither of these alternatives is viable in the absence of an accurate chromatic vision diagnosis.

4.2. Detecting and Diagnosing Color Blindness: A Two-Step Strategy

Pseudoisochromatic tests (Fig. 4) use stimuli that are very similar in color ("isochromatic") for color-blind people but not for normal-sighted people ("pseudo"). For example, in the famous Ishihara test (Fig. 4A), color-blind people have difficulty in reading the number presented in each plate. This is because they do not perceive enough differences between the points that form the number and the background. Another similar
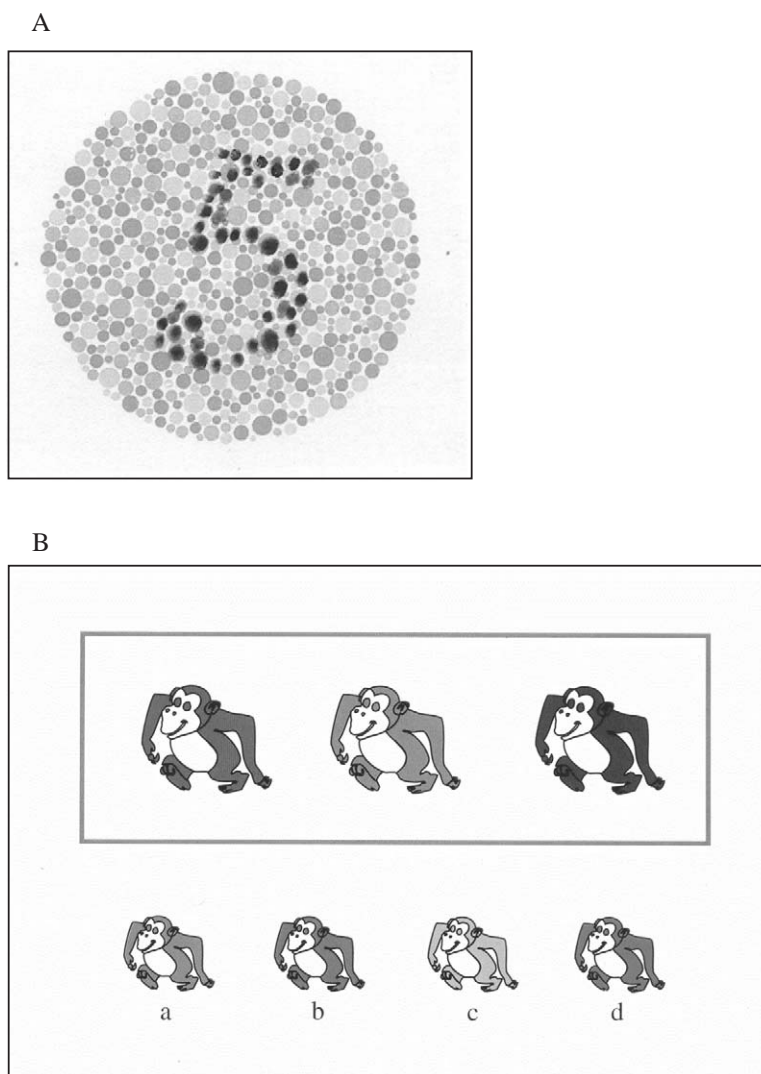
FIGURE 4 Two examples of pseudoisochromatic plates (originals are colored). (A) Ishihara test (observer's task is to identify the number). (B) TIDA test (observer's task is to select the little monkeys similar in color to the big monkeys).

example is provided by a test, called the "TIDA," to detect color-blind children. In this case (Fig. 4B), color-blind individuals of the red–green type select some greenish and reddish monkeys from the bottom line as being similar ("belonging to the same family") to the larger gray monkeys presented in the upper line.

Because they are easily constructed and applied, there are many pseudoisochromatic tests available for using during the first stage of the diagnostic process to detect the presence of color blindness. However, their use must be complemented, during a second stage, with an anomaloscope to get the most accurate

diagnosis. Among the pseudoisochromatic tests capable of detecting a red–green anomaly are the following: the Ishihara test; the Dvorine test; the Hardy, Rand, and Rittler (HRR test); the D-15 test; and the City University Color Vision Test. The latter three tests are also useful for detecting tritan anomalies and for grading the severity of the color perception problem.

Anomaloscopes show users a circle divided into two halves. One half is used to present the reference stimulus and must be metamerized in the other half by a stimulus mixture. As indicated previously, to detect and differentiate red–green problems, a Nagel match must be used.

This kind of match uses a reference stimulus of 589 nm (yellow for a normal-sighted person) and an adjustable stimulus mixture made up of 546 nm (green for a normal-sighted person) and 670 nm (red for a normal-sighted person). To detect and differentiate among tritan problems, the Moreland match must be used. This uses a reference stimulus made up of a mixture of 480 nm (greenish blue for a normal-sighted person) and 580 nm (yellow for a normal-sighted person) and an adjustable stimulus mixture made up of 436 nm (blue for a normal-sighted person) and 490 nm (bluish green for a normal-sighted person).

4.3. After Diagnosis: Trying to Compensate for Perceptual Limitations

Gaining awareness of people's color blindness and perceptual limitations is the first step in the process of obtaining psychological help. With this knowledge, color-blind people can avoid incorrect attributions about the causes of their perceptual difficulties, and just as important, they obtain a first cue to adopt compensatory interventions. More concretely, and returning to the case discussed earlier, Mr. X's awareness of his perceptual peculiarities prevented him from attributing his difficulty in locating red marks to cognitive or motivational factors. He never thought that his difficulties were due to an attentional deficit. Eventually, he realized that his problems were the consequence of his difficulty in differentiating the dark red ink marks from the black letters, and so he was able to suggest the use of blue ink.

Knowing that a person is color-blind, the psychologist should look for, and try to compensate for, the following two kinds of problems:

- Difficulty in performing figure–background segmentations
- Difficulty in using the basic chromatic categories

4.3.1. Figure–Background Segmentation Difficulties

Many computer users are familiar with the following situation. Initially, they used a common application (e.g., Microsoft PowerPoint) to present a colored written message on a colored background (e.g., yellow letters on a green background). Of course, this message is easy to read on the original screen because the letters used have an adequate lightness contrast with the background.

However, when another screen is used to present the same message (e.g., a large screen used in a classroom), the message becomes very difficult to read due to a reduction in the lightness contrast. The cause of this problem is simply that the same document, as well as the same software specification, produces different colors on different screens.

A problem very similar to the one just described is relatively frequent for some types of color-blind people because, when looking at some stimuli, they experience lightness contrasts that are different from those experienced by normal-sighted people. For example, protanopes perceive reddish backgrounds as darker and, consequently, have difficulty in reading black letters on reddish backgrounds. They also have difficulty in detecting when a red light is on because they perceive it as less brilliant than do normal-sighted people.

Color-blind people not only have figure–background segmentation problems derived from their peculiar lightness perception, but a similar type of problems also may be a consequence of their restricted gamut of chromatic experiences. To understand the origin of this type of difficulty, consult Fig. 3 again and remember that, for some dichromats, the three stimuli used by the anomaloscope affect the responses of only one cone type. Consequently, when presented with the precise physical intensity (see Eq. 1), the exact same color experience is produced in these people, although normal-sighted people would perceive very different hues (green, yellow, and green) in response to these stimuli.

4.3.2. Problems Related to the Use of Basic Chromatic Categories

The expression "basic chromatic categories" is used to name the linguistic categories consistently used among speakers of a certain language. For the languages in most developed countries (e.g., English, Spanish, Japanese), there are 11 categories. These are identified in English by the following terms: red, green, blue, yellow, white, black, brown, orange, purple, pink, and gray. To say that these categories are used consistently means that, for example, if an observer is required to select the prototype of one of these categories, he or she will select the same one, or a very similar one, on different occasions. These choices will also be very similar to the equivalent choices performed by other normal-sighted people.

Previously, the singular adjustments performed by color-blind people in a Nagel anomaloscope were described. Specifically, it was indicated that stimuli

belonging, for normal-sighted people, to the red, green, and yellow categories can be metamers for red–green dichromats. Taking this into account, and also considering that color-blind people generally have an increased number of metamers, it is logical to expect color-blind people to have severe difficulties when using basic chromatic categories. In contrast to this expectation, research carried out with 30 dichromat children (protanopes and deuteranopes) provided the results presented in Fig. 5. The children achieved an important level of "hits" when required to name prototypical exemplars of the 11 basic Spanish chromatic categories. Moreover, for five basic categories (yellow, green, blue, red, and white), their hit percentage was more than 85%.

Why is there such disparity between, on the one hand, the low discriminative level shown by color-blind people in response to the stimuli used by diagnostic tests and, on the other, their relatively high capacity to name prototypes? What is the practical relevance of this fact? These questions are addressed in turn.

The main difference between the stimuli used in an anomaloscope and the stimuli provided by everyday chromatic surfaces is the range of available intensities. For example, when using a Nagel anomaloscope, the wavelengths that a normal-sighted person perceives as green (546 nm), yellow (589), and red (670 nm) can be presented with a broad range of physical intensities. Consequently, differences in the capacity to activate

cone responses can be compensated by intensity variations (see Eq. 1). This compensation is not possible with surface colors in which the response intensity (perceived lightness) is very closely related to the specific wavelengths reflected. So, for red–green dichromats, the yellow prototype is always perceived with a higher lightness than is the green or red prototype. Of course, this also helps these individuals to avoid naming confusions. This does not occur when no prototypical exemplars of the basic categories are used (e.g., pale green), in response to which more naming errors are produced.

Because many everyday chromatic stimuli are close to the basic chromatic category prototypes, color-blind people can frequently name a color correctly (Fig. 5). Consequently, for example, many teachers frequently do not agree with the results provided by the standardized chromatic vision evaluation procedure because these results are always worse than children's real capacity to name and differentiate everyday chromatic materials. Taking this into consideration, should we conclude that chromatic evaluation is useless? The answer is clearly no because color blindness detection alerts one to the problems with figure–background segregation discussed previously and, logically, because these individuals continue to have more trouble in naming and using color categories than do normal-sighted people, even when prototype exemplars are used (Fig. 5).
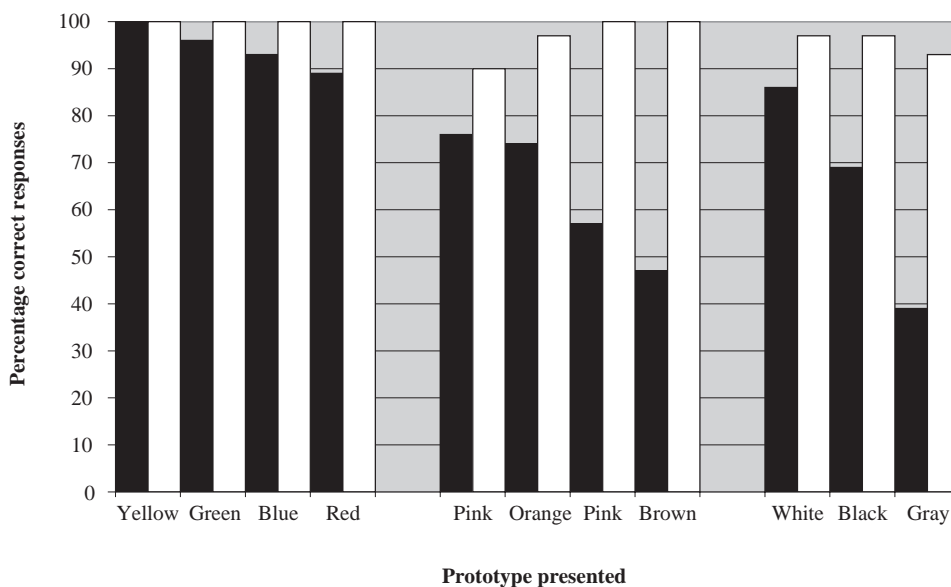

FIGURE 5 Percentage of correct naming responses to basic Spanish category prototypes. Black bars, dichromats' responses; white bars, controls' responses.

See Also the Following Articles

Coping ■ Psychoneuroimmunology

Further Reading

Birch, J. (1993). *Diagnosis of defective color vision.* Oxford, UK: Oxford Medical Publications.

Fletcher, R., & Voke, J. (1985). *Defective color vision: Fundamentals, diagnosis, and management.* Boston: Adam Hilger.

Hsia, Y., & Graham, C. H. (1997). Color blindness. In A. Byrne, & D. R. Hilbert (Eds.), *Readings on color: The science of color.* Cambridge, MA: MIT Press. (Original work published 1965).

Kaiser, P. K., & Boynton, R. M. (1996). *Human color vision* (2nd ed.). Washington, DC: Optical Society of America.

Mollon, J. D., Pokorny, J., & Knoblauch, K. (2003). *Normal and defective colour vision.* Oxford, UK: Oxford University Press.

Newton, I. (1952). *Optics* (4th ed.). New York: Dover. (Originally published 1730).

Community Psychology

Kenneth I. Maton

University of Maryland, Baltimore County, Baltimore, Maryland, USA

GLOSSARY

ecological In psychology, an emphasis on the importance of the environmental context in understanding individual behavior and social problems.

prevention Interventions that aim to reduce the occurrence of physical or psychological disorders and social problems.

promotion Interventions that aim to enhance psychological well-being and the quality of community life.

social action Interventions that strategically exert influence at the grassroots level to enhance access to resources for disadvantaged groups.

Community psychology is concerned with the relationship between individuals and their social and community environments. Community psychologists contribute to the public welfare through collaborative interventions that prevent psychological and social problems, promote personal and community well-being, and empower disadvantaged citizens and groups.

1. INTRODUCTION

The field of community psychology developed during the mid-1960s as a reaction against the limitations of traditional intrapsychological approaches to research and social problems. In contrast, the developers of the field emphasized the importance of social contexts and social environments in understanding behavior and social problems, envisioned a strengths-based approach rather than a deficits-based approach to research and action, and emphasized the importance of the prevention of problems rather than the treatment of problems after the fact. Community psychology shares with sociology an emphasis on the influence of structural features of the environment on human behavior, but the former differs from the latter research discipline in its greater emphasis on action and individual experience. Community psychology shares with social work a focus on helping individuals in a community context, but the former differs from the latter applied field of practice in its greater emphasis on empirical research, systems-level change, and prevention.

Community psychologists are involved in research and action focused on the entire spectrum of contemporary issues facing citizens and communities. These include educational reform, homelessness, domestic violence, intergroup conflict, substance abuse, teenage pregnancy, poverty, child abuse, youth violence, racial and sexual discrimination, coping with life stress, community capacity building, and enhancing psychological sense of community. Applied research on these issues draws on a wide

range of research methods, including needs assessment, qualitative interviews, quantitative field research, focus groups, evaluation research, and action research. Similarly, the range of intervention approaches through which community psychologists address social problems is broad and includes program development, consultation, community coalition building, social action, community development, and social policy advocacy.

Distinctive aspects of community psychology include its guiding values, its ecological multilevel framework, and its focus on prevention, promotion of well-being, group empowerment, and community change. Each of these is described in this article.

2. GUIDING VALUES

Community psychologists emphasize the distinctive nature of the values that guide their work. The following seven guiding values can be discerned.

2.1. Individual Wellness: The Physical, Psychological, Social, and Spiritual Health of Citizens

Unlike clinical psychologists, community psychologists do not focus on the etiology and remediation of mental illness; rather, they focus on understanding the factors that contribute to well-being and healthy development and on the associated development of programs that prevent the development of physical and psychological problems.

2.2. Sense of Community: Citizens' Sense of Belonging to a Larger Group

Community psychologists view the decline of community cohesion as a major contributor to individual and social problems and the enhancement of a sense of community as a key contributor to individual wellness and community vitality.

2.3. Social Justice and Empowerment: The Equitable Distribution of Economic, Political, and Psychological Resources in Society

Community psychologists are concerned with addressing the pressing social issues of the day and in particular with empowering marginalized groups through enhancing their access to economic, political, and psychological resources.

2.4. Citizen Participation: The Active Involvement of Citizens in All Aspects of Community Life

Community psychologists are wary of expert-driven, top-down solutions for the problems that communities face, viewing such approaches as too often ineffective and as contributors to citizen alienation and dependence. In contrast, community psychologists view citizen involvement as the antidote necessary to bring about community revitalization and solutions to pressing social problems.

2.5. Collaboration and Community Strengths: Working Together with Citizens and Groups in the Community and Building on Their Strengths

As proponents of a strengths-based perspective, community psychologists view citizens and community groups as the sources of solutions to community problems. Working together as partners allows the respective strengths of both community members and community psychologists to contribute to community betterment.

2.6. Respect for Human Diversity: Genuine Appreciation of Human Differences Such as Differences in Ethnicity, Gender, Sexual Orientation, Religion, Age, Physical Disability, and Social Class

Community psychologists are critical of social paradigms that view those who are "different" from the mainstream as the cause of the problems they face in society. In direct contrast, community psychologists affirm diversity as a public good and an invaluable resource.

2.7. Empirical Grounding: The Research Basis for Efforts to Make a Positive Difference in Communities and the Larger Society

Community psychologists view systematic evidence as the key to understanding individual and social problems and view such understanding as central to

effective program development, community action, and social policy.

Taken as a whole, these seven values reflect the field's joint commitment to research and action, its focus both on bettering the lives of individuals and on bringing about broader social change, and its emphasis on working with, for, and in the community. Individual community psychologists, however, vary greatly on the relative importance placed on each of the guiding values, with resultant differences in preferred approaches to community work. For example, academic researchers emphasize the importance of research and publication, whereas applied community practitioners emphasize the value of action and hands-on community involvement. Community psychologists oriented to individual wellness emphasize the development of local programs to prevent problems and assist those in need, whereas those oriented to social justice and empowerment often articulate the need for higher level social change efforts to transform social norms and to redistribute social and economic resources. Strongly quantitative community psychology researchers emphasize traditional research values of rigorous design and measurement validity, whereas those more oriented to social justice issues emphasize discovery-oriented qualitative, participatory, and action research methods.

3. ECOLOGICAL MULTILEVEL FRAMEWORK

The guiding conceptual framework of community psychology is social ecology, which emphasizes that an understanding of human behavior must always encompass the social environment in which it is embedded. Within the social environment, multiple levels of analysis can be distinguished, including the individual, the microsystem (e.g., family, classroom, social network), the organization, the community, and the macrosystem (e.g., cultural norms; the larger society). Research at the individual level of analysis might focus, for example, on the relationship between psychological well-being and level of perceived social support in a sample of inner-city youth. At the microsystem level of analysis, a researcher might study the relationship between the social climate of self-help groups and group effectiveness. Evaluating the impact of a comprehensive, community-wide teenage pregnancy prevention program on incidence rates of teenage pregnancy in communities that either received or did not receive the

intervention is an example of research at the community level of analysis. Additional examples of research topics at various levels of analysis are depicted in Table I.

Central to the ecological paradigm is the principle of ecological interrelatedness. Lower levels of analysis are embedded within higher levels and are directly influenced by, and can also influence, these higher levels. Thus, the role of the environment in causing individual and social problems is emphasized. Indeed, community psychology arose in part as a reaction against views that "blamed the victims," such as lower income populations and ethnic minorities, for problems that community psychologists view as stemming from larger community and societal forces.

Indeed, given the complex social problems with which the field is concerned, community psychologists believe that multiple levels of analysis should ideally be encompassed in community research and action. Take, for example, the case of youth violence in urban areas. A community psychology approach to addressing that problem, based on the ecological perspective, would naturally call for interventions at multiple levels. These might include interpersonal problem-solving and

TABLE I
Ecological Levels of Analysis in Community Research

Level	Sample research topics
Individual	Psychosocial predictors of well-being, competency-enhancing prevention programs, psychological empowerment, resilience, individual indicators of community problems
Microsystem	Family cohesion, social networks, peer groups, classroom social climate, teacher expectations, self-help group processes, neighborhood effects
Organization	Organizational characteristics of empowering community settings, whole school reform, faith-based organizations, voluntary associations, block associations
Community	Community coalitions for drug abuse prevention, community-wide health promotion, psychological sense of community, community readiness for change
Macrosystem	Social–political forces, cultural beliefs, social movements, strengths-based social policy

Adapted from Dalton, Elias, & Wandersman (Table 1.1, p. 13).

conflict resolution skills training at the individual level, family support programs at the microsystem level, whole school reform of urban schools at the organizational level, a campaign to alter cultural norms about violence and masculinity at the community level, and a strengths-based antipoverty initiative at the federal level. Although a given research or intervention project of any individual community psychologist will not realistically be able to encompass all ecological levels in any social problem area, the field as a whole aspires to conduct research and interventions that encompass the entire spectrum of ecological levels.

4. COMMUNITY RESEARCH METHODS

Community psychology researchers employ a wide range of research methods to examine community phenomena of interest. These encompass both qualitative and quantitative methods as well as research designs ranging from experimental program evaluation to action research.

Qualitative methods are especially useful for generating in-depth descriptions of a phenomenon, an understanding of context, new insights, and opportunities for the "voice" of marginalized populations to be heard. Methods include qualitative interviewing, participant observation, and focus groups. Interviews can be coded systematically using sophisticated qualitative software programs. Participant observation allows in-depth examination of a particular group or community setting, with a focus on uncovering the worldview and meaning system of members. Focus groups facilitate the generation of common themes that emerge as a result of discussion in the group context.

Quantitative methods are used for varied purposes in community psychology, including hypothesis testing, needs assessment, and program evaluation. Hypothesis testing research may focus on community samples of individuals, on samples of groups or organizations, or (increasingly in multilevel research) on individuals, families, and groups nested in a sample of organizations, neighborhoods, or communities. Needs assessment provides systematic descriptive information concerning the nature and level of local needs in a setting or community and serves as a basis for collaborative planning of interventions. Program evaluation examines the extent to which interventions are achieving their intended outcomes and the factors contributing to observed outcomes, using experimental or quasi-experimental research designs.

Increasingly, community researchers combine qualitative and quantitative methods within a single study. This approach allows for building on the strengths of each research method. Furthermore, to address community phenomena, there is increasing focus on interdisciplinary research that incorporates research methods from other fields. Examples include community ethnography (anthropology), geographic information systems (geography), community case study (community sociology), and epidemiology (public health).

Consistent with the community psychology values of citizen participation and collaboration as well as building on community strengths, partnerships are often developed between community researchers and populations under study. Investigators engage those being studied as active collaborators in defining the research question, helping to develop appropriate research instruments, designing and implementing research procedures, and interpreting and disseminating research findings. In some cases, helping local groups or organizations to develop their own research capacities may be one of the intended outcomes of the research process, as is the case in empowerment evaluation.

Community psychology espouses adventuresome research methods appropriate to the research question at hand, even if these methods are not conventional in nature. In many research contexts, the researcher is viewed not as an objective, fully detached observer but rather as an individual whose values influence the scientific enterprise in important ways. Furthermore, consistent with this perspective, the researcher's relationship with those studied is seen as part of the larger social ecology of the phenomenon under investigation.

5. COMMUNITY RESEARCH AREAS

Community psychologists examine a wide range of topics in community-based research. Four illustrative research areas are briefly described in what follows.

5.1. Social Problems

One broad area of community research centers on pressing social problems such as discrimination, poverty, education of minority students, substance use, HIV/AIDS, and homelessness. Much of this research focuses on marginalized groups in society such as ethnic minorities and lower income populations. Community

psychologists examine factors that contribute to the specific social problem under focus. Individual, family, and environmental variables are encompassed in this research. For example, in the urban education area, community psychology researchers have examined the roles of family support, teacher expectations, the transition to middle school, the social climate and culture of schools, and neighborhood poverty. The results of research in each social problem area provide a basis for intervention approaches to address the particular social problem.

5.2. Stress and Mental Health

A second long-standing focus of community psychology research examines the interrelationships among stress, coping, social support, and psychological well-being. Life stressors examined run the gamut from acute stressors (e.g., death of a parent) to ongoing stressors (e.g., poverty). Research in this area has revealed various personal and environmental factors that help individuals to cope with stress, including spirituality, interpersonal problem-solving skills, and social support from friends, family, and mentors. This body of research has suggested important foci for psychological and environmental interventions that aim to prevent negative outcomes of stress and enhance resilience in the face of adversity.

5.3. Community and Quality of Life

A third area of research examines the role of community settings in contributing to the quality of individual and community life. This research encompasses settings such as voluntary associations, block associations, faith-based organizations, and self-help groups. These settings provide a meaningful niche in which citizens find meaning, support, and opportunities to contribute to the well-being of others and the larger community. Research on community settings indicates that certain organizational characteristics, such as cohesion, support systems, opportunity role structure, and leadership, have been found to be related to member well-being and psychological empowerment.

5.4. Sense of Community

A fourth illustrative area of community psychology research centers on psychological sense of community. This refers to the sense of belonging or connectedness to a larger group. Sense of community is important both in relational communities defined by shared goals and activities (e.g., church, professional association) and in geographical communities (e.g., neighborhood, city, nation). Researchers have delineated four underlying dimensions to sense of community: membership, influence, integration and fulfillment of needs, and shared emotional connection. The individual and environmental correlates of sense of community, subgroup differences, and changes over time represent important topics of research in this area.

The importance of context and the importance of diversity are cross-cutting themes in many areas of community psychology research. Depending on the dangerousness of neighborhood context, for example, parenting style or psychological sense of community has been found to be more or less adaptive. Ethnic groups differ in terms of life contexts and challenges, culturally embedded values and meaning systems, and access to resources and services. Findings related to distinctiveness of context and diversity provide a critical foundation for the development of context-sensitive and culturally competent interventions.

In addition to theory generation and hypothesis testing research, many community psychologists conduct evaluation research. This research sometimes focuses on interventions developed by the investigator but more frequently examines various community-based programs developed by others. Evaluation research examines program outcomes, the quality of implementation of programs, and the theory-based processes through which program efforts occur.

Importantly, research in community psychology often builds on theory and findings from other subfields of psychology. Developmental psychology provides critical underpinnings for community psychology research on children's coping, support, and well-being. Social psychological theory contributes to community psychology work in the areas of discrimination, oppression, and empowerment. Health psychology models help to guide community psychology work in the areas of health promotion and disease prevention. Cultural psychology provides important theoretical models and perspectives for community psychology work on ethnic identity, coping, and ethnic minority populations.

In comparable fashion, disciplines outside of psychology influence community psychology theory and research. Public health has a major influence in the area of prevention and health promotion. Anthropology's focus on cultural norms, practices, and worldviews has broadened community psychology models to encompass the culture of settings and cultural narratives of community members. More generally, across the gamut of social

problem domains, there are major theoretical paradigms and accumulated research findings from other fields of direct relevance (e.g., sociology of education, community sociology, urban affairs, criminology, economics, law, social history).

Conversely, many community psychology models and perspectives have influenced other subareas of psychology and allied researchers in other disciplines. For example, prevention has become an important domain of investigation in other applied areas of psychology, including developmental psychology, applied social psychology, health psychology, and clinical psychology. Similarly, the importance of community-based research, ecological context, the strengths-based paradigm, qualitative research, and varied community psychology findings in specific social problem areas has influenced investigators in multiple areas of psychology as well as those in allied social science disciplines.

6. PREVENTION AND PROMOTION

In contrast to helping individuals after serious problems have developed, community psychologists seek to prevent the initial occurrence of problems. Primary preventive interventions aim to reduce the number of new cases (i.e., incidence) of a wide range of physical illnesses and psychological problems, including HIV/AIDS, substance abuse, and depression. Universal preventive interventions seek to do so while focusing broadly on a general population (e.g., all students in a school or school system). Selective preventive interventions focus strategically on the subset of individuals at above average risk for developing a disorder or problem due to either environmental factors (e.g., alcoholic parent, concentrated poverty) or personal factors (e.g., school difficulty). Indicated prevention programs focus specifically on individuals at high risk for developing a disorder or problem, usually those showing early symptoms of the disorder or problem.

Prevention programs have been developed for every stage of the life span. However, most prevention programs focus on infants, children, and youth, based on the principle that early intervention is the most likely means to prevent later programs. For example, Nurse Home Visitor programs provide low-income, single new mothers with support, child-rearing information and skills, and linkages to community-based resources as a means to prevent child abuse and poor child outcomes. Head Start and related preschool problems help low-income preschoolers to learn basic school-related social and cognitive skills commonly present in their middle-class counterparts, thereby preparing those served to make a positive transition to formal schooling.

Prevention programs in public schools provide at-risk children with important coping skills and social support. Programs for middle and high school youth often focus on prevention of specific social problems such as substance abuse, school dropout, and teenage pregnancy, for example, by providing refusal skills to counteract negative peer pressure, information that underscores the implications of poor choices, and meaningful positive activities (e.g., community service) in which to engage.

Prevention specialists increasingly emphasize the importance of comprehensive, community-wide, and multilevel interventions. In the case of a program to prevent school dropout among urban youth, for example, a comprehensive preventive intervention might address multiple risk factors related to dropout, such as students' academic skills and knowledge, coping skills, engagement in creative or prosocial activities, and adult social support, rather than only one of these (and other) contributing factors. A multilevel approach in the academic performance domain, for example, might include an individual tutoring component, enhancement of the cultural competence of teachers, and schoolwide changes to provide a more supportive and caring school climate rather than an intervention at only one of these levels. Finally, a community-wide intervention could mobilize students, teachers, parents, nonprofit organizations, businesses, and the media. Given the difficulty of making a difference in pressing social problems, comprehensive, multilevel, and community-wide programs are more likely to provide the intensity and multiplicity of impact necessary to counter the multiple factors contributing to problem development.

Promotion of wellness interventions seeks to enhance psychosocial competencies and development, as well as the overall quality of individual and community life, rather than aim specifically (as preventive interventions do) at reducing rates of a specific disorder or problem. Intervention programs that contribute to normative child development, enhance children's resilience to adverse circumstances, and/or strengthen parenting skills are examples of promotion interventions at the individual level of analysis. Efforts to enhance psychological sense of community, intergroup relationships, and/or neighborhood revitalization are examples of promotion interventions at the microsystem or community level.

Outcome research has demonstrated the effectiveness of a range of prevention and promotion programs. However, these programs often are developed by university-based investigators under the ideal conditions of large amounts of grant funding and a singular focus on program success. A major challenge in the prevention area has been to effectively sustain programs after grant funding ends and to effectively disseminate programs to new sites and communities lacking the resources and special conditions that were present for initial program development. Given these challenges, the study of factors leading to effective program implementation, dissemination, and adoption is receiving additional focus. Some of the key variables found to be important include strong agreement on program goals among the various stakeholders, linkage of the program goals to the basic mission of the setting, clear and strong program leadership, and adequate implementation of core program components and principles.

7. SOCIAL ACTION AND COMMUNITY CHANGE

An additional important domain of community psychology interventions relates to social justice and empowerment. Social action and community change aim to enhance levels of economic, social, and psychological resources for individuals and groups lacking power in society. A range of methods are employed.

Grassroots organizing efforts bring together citizens and community groups in low-income neighborhoods to pressure elected officials to allocate additional funding for basic individual and community needs. Organizing efforts may involve recruiting individual citizens or working with extant community groups, such as inner-city churches, that have preexisting access to large numbers of citizens. Community psychologists involved in grassroots organizing efforts bring to bear their group skills and empowerment perspective along with linkages to resources and research bases that can be used by those leading the organizing efforts.

Community coalition development brings together groups from multiple sectors of the community to devise plans for addressing a specific community need. Groups represented often include local government, business, human service agencies, citizen associations, and faith-based organizations. Community

psychologists can serve as coalition conveners, resource experts in specific substantive areas, technical support providers, and/or evaluators.

Action research addresses local community problems through collaboration with community groups centered on local community needs. Projects may be initiated by community psychologists interested in a particular community problem or by community groups that approach community psychologists for help in addressing a need specific to their groups or the community. Consistent with an empowerment perspective, the research and action pursued have direct utility for self-defined local community needs.

Finally, community psychologists have become increasingly involved in social policy advocacy. Some community psychologists bring policy-relevant research to the attention of policymakers through congressional hearings, dissemination of reports, and/or congressional briefings. Others work full-time in the policy sector as congressional aides or policy analysts in the executive branch of government. Finally, a small number of community psychologists have become policymakers in elected positions, including local school board members and city council members.

In their social action and community change work, community psychologists necessarily work with a wide range of citizens, community groups, organizational staff and directors, and elected officials. The values of the field, including the focus on empowerment, collaboration, and community strengths, serve as important guides in this arena.

See Also the Following Articles

Social Networks ■ Stress

Further Reading

Dalton, J. H., Elias, M. J., & Wandersman, A. (2001). *Community psychology: Linking individuals and communities*. Belmont, CA: Wadsworth/Thompson.

Levine, M., & Perkins, D. V. (1997). *Principles of community psychology: Perspectives and applications* (2nd ed.). New York: Oxford University Press.

Rappaport, J., & Seidman, E. (Eds.). (2000). *Handbook of community psychology*. New York: Plenum.

Revenson, T. A., D'Augelli, A. R., Hughes, D., Livert, D., French, S. E., Seidman, E., Shinn, M., & Yoshikawa, H. (Eds.). (2001). *A quarter-century of community psychology*: New York: Plenum.

Rudkin, J. K. (2003). *Community psychology: Guiding principles and orienting concepts.* Upper Saddle River, NJ: Prentice Hall.

Sarason, S. B. (1974). *The psychological sense of community: Prospects for a community psychology.* San Francisco: Jossey–Bass.

Shinn, M., & Toohey, S. M. (2003). Community contexts of human welfare. *Annual Review of Psychology, 54,* 427–459.

Weissberg, R. P., & Kumpfer, K. L. (2003). Prevention that works for children and youth [special issue]. *American Psychologist, 58*(6/7), 425–490.

Compensation

Henk Thierry

Tilburg University, Tilburg, The Netherlands

GLOSSARY

cafeteria plan A system in which organization members get the opportunity to trade some labor conditions for others to design a package with the highest personal utility; work time may be traded for leisure time, cash may be traded for additional health insurance, and so forth.

fringe benefits A term that refers to secondary and tertiary labor conditions; primary conditions cover direct compensation and provisions on work time, holidays, and vacations, secondary conditions concern health insurance and other types of social security, and tertiary conditions are specific to a company (e.g., car lease rates).

incidental pay A system that provides a supervisor with the opportunity to rapidly express recognition for a job extremely well done, for effort beyond the call of duty, and the like; rewards are often small (e.g., a free lunch, a dinner for two, a paid conference).

job evaluation A systematic procedure for describing the standard content of a job as well its standard conditions; depending on the system's particular characteristics, job descriptions are analyzed and evaluated.

merit rating A system of performance appraisal based on qualitative characteristics of individual workers (e.g., creativity, diligence, care for clients); the appraisal outcome is rewarded (usually annually) with an extra step on the job's salary scale or with a separate bonus.

performance-related pay A category of pay systems in which work outcomes are tied to pay; outcomes may bear on individual or team contributions to work, the achieved results, and/or departmental or organizational effectiveness indexes (e.g., sales volume, profits, customer satisfaction).

reflection theory on compensation A theory in which compensation is considered to be a major domain relevant to a person's self-identity; compensation reflects information in four areas and resulting in four different meanings: motivational properties, relative position, control, and spending.

Compensation is a major domain within human resource management to which several disciplines contribute. Psychological approaches to financial compensation focus on the motivational force of pay, its potential to get organization members to learn new abilities and skills, and the conditions for perceiving pay systems and practices as fair. In industrialized countries, the structure of compensation in profit and nonprofit organizations is quite comparable. Two important components of compensation—job evaluation (which often determines base pay) and performance-related pay—are considered more closely. Research indicates, for instance, that job evaluation tends to grasp one common factor (extent of required education), whereas employees may perceive the evaluation outcome as a harness. Performance-related pay appears to be a rather vulnerable kind of system, particularly when performance appraisal is based on qualitative characteristics. Frequently, employees do not

perceive a relationship between the appraisal result and the awarded bonus amount. When quantified critical performance factors are applied, this relationship becomes more clear. Also, rather small differences in pay affect the behavior at work more often than do large differences in pay.

1. PSYCHOLOGICAL PERSPECTIVES

Managers and employees in an organization often wonder whether its compensation systems and strategies affect the work behaviors of the organization's members, individually as well as in teams, and the performance results of the organization at large. Psychologically, this theme can be addressed in more than one way. One approach is in terms of the conditions that make compensation more or less motivating, often relative to the motivating potential of other work characteristics. Characteristically, the motivational force of compensation is inferred from the organization members' level of effort, achieved performance, degree of satisfaction, and opportunities for personal development as well as from the effectiveness of cooperative efforts. A slightly different approach concerns the extent to which compensation facilitates the learning of new abilities and skills rather than stimulating employees to maintain their former habitual working patterns. This second approach is also basic to the decision, in cases of complex processes of organization change, as to whether alterations in compensation should precede or follow these. A third approach is distinguishable as managers face, for instance, the need to cut costs and are considering whether another compensation package might help them to achieve this. In such cases, an important perspective is the fairness with which this may be brought about, both in the procedures followed and in the quality of the end result. Here again, it is vital to consider the conditions under which compensation affects performance results.

These (and other) psychological perspectives are becoming more meaningful as they are considered within the frameworks of particular compensation systems. Although there are huge differences among pay forms and systems, there is a comparable compensation design across industrialized countries. This design is outlined in the next section. Then, two main categories of systems—job evaluation and pay for performance—are discussed from the viewpoint of how psychological insights may help to better understand recurring problems.

2. THE DESIGN OF COMPENSATION

Although there are great differences among industrialized countries in the amount of pay provided for particular jobs, the structure of compensation is quite comparable throughout the major part of the world. It is designed on the basis of most or all of the components described in what follows.

2.1. Base Pay

The base pay component reflects the value of a job to an organization often with respect to the particular labor market of a country or region. In many cases, systems of job evaluation are used to appraise and order the value of most or all available jobs within an organization. These relative job values are subsequently categorized in classes to which particular wage and salary scales are tied. In other cases, managers monitor the salaries that are paid in referent organizations for particular jobs and apply these to achieve "market conformity." Job values embody, to a large extent, how scarce particular abilities and skills—or competencies—are within a (national) labor market. Thus, base pay recognizes the (scarcity of) qualifications of a worker to fulfill a specific set of tasks.

2.2. Performance-Related Pay

The performance-related pay component bears on the appraisal of job performance. Characteristically, specific targets or standards are set, and these have to be met or passed before a particular bonus is awarded. When the quantity of work performance was dominant, targets applied to the number of "pieces" made (piece rate system) or the amount of time required to perform a particular task (tariff system). Current standards relate more often to performance quality, team effectiveness, sales value, client satisfaction, and the like. Through performance-related pay, employees are recognized for what they do (and don't do) in their work and are often thought to be stimulated to improve their performance.

2.3. Fringe Benefits

The fringe benefits component reflects the labor conditions that apply to members of an organization. In Europe, fringe benefits usually relate to a collective labor agreement (within the framework of a country's social security policy) and include both secondary (e.g., health insurance programs) and tertiary conditions (e.g., a

company's car lease program). In the United States, fringe benefits packages are much more organization specific, mirroring entrepreneurial and managerial views on appropriate provisions. Usually, their coverage is much more modest than in European countries. When a cafeteria plan is applied, organization members get the opportunity to select, on a regular basis (often annually), alternative labor conditions, trading less attractive options for more attractive ones. Fringe benefits mirror the views of entrepreneurs and union leaders (and society as well) on the required level of minimal security in work and life.

2.4. Incidental Pay

When deciding on matters of pay, managers in most countries must take into account various laws, regulations, and the specific organization's tradition of established practices. This host of rules might make it very hard to quickly express recognition to employees who have performed beyond expectations, put in many more hours than agreed on to speed up deliveries, and so forth. Incidental pay provides managers with the opportunity to reward some employees directly when applicable. Usually, the awards are small (e.g., a lunch for two, a dinner party). However, incidental pay might also involve larger awards (e.g., a paid conference).

Countries differ to a large extent regarding the proportional amount of base pay within their organizations. In some countries, base pay may amount to more than 90% of an employee's salary, whereas in other countries, base pay may cover approximately 50% of the salary. Moreover, opinions vary concerning the degree to which organization members' incomes should be based on their performance and sensitive to the volatility of (consumer) markets. Because job evaluation (base pay) and pay for performance have an incisive effect on managers' and employees' behaviors at work, these two components are addressed in more detail in what follows.

3. JOB EVALUATION

There are many different systems of job evaluation in use. Common to most systems is that key jobs are described according to their content and accompanying standard work conditions and are subsequently analyzed and appraised in terms of a job evaluation system's characteristics (e.g., required knowledge, problem solving, supervision). Frequently, a point rating procedure, in which numerical values are assigned to the main job features, is applied. The sum total of these values constitutes the "job value." This approach is, to some extent, derived from the technique of job analysis, a well-known subject area within work psychology. Through job analysis, the work elements of a job (e.g., cutting; sewing) are carefully analyzed in terms of those worker characteristics that are required for successful job performance (e.g., abilities, skills, personality factors). The objective of any job analysis system is to grasp the essence of a job in terms of a variety of distinct "worker" and/or "work" qualifications.

Would the latter objective also hold for job evaluation? Research shows that these systems usually measure one common factor: extent of required education. In practice, several (professional) groups in organizations tend to stress the relevance of particular job qualities (e.g., leadership behavior, manual dexterity and sensitivity for specific tools in production work) that are believed to have been underrated in a particular job evaluation system. But factor-analytic studies show that even when a job evaluation system is "extended" to include such particular job qualities, second and/or third factors—if they can be identified at all—have much in common with the first (general education) factor. In other words: job evaluation measures the educational background required for adequate decision making in a job. Although the concepts and wording of job characteristics in various job evaluation systems might seem to differ from one another, they tap into the same construct. This is also exemplified through the custom, in several countries, of using a simple arithmetic formula to translate job value scores from one system of job evaluation into the scores of another system.

Obviously, these procedures for determining job value and base pay level do have important implications. First, the process of translating work descriptions or characteristics into worker attributes is quite vulnerable because it requires psychological expert knowledge to determine the variety of attributes (e.g., abilities, aptitudes, skills) relevant to adequately perform a particular job. Research has shown repeatedly that ratings may be biased in various respects due to the "halo effect," gender discrimination, implicit personality theory, and the like. This theory refers to a pattern among attributes that a particular rater believes exists. One example might be that somebody scoring high on emotional distance is thought to excel in analytical thinking, have a "helicopter view," and engage in a structuring leadership style because the rater assumes such a pattern to be "logical."

Second, regardless of an individual worker's personal conception of his or her job's main content, that job's

value is determined mainly by the level and nature of the required educational background. Consequently, a worker may perceive the evaluation of his or her job as not representing the worker's ideas about the job's actual content but rather as being the outcome of a bureaucratic exercise. Related to this is a third issue: Job evaluation may be experienced as a "harness" (because it measures one factor nearly exclusively), especially when a fine-grained salary structure is tied to the job value structure. This characteristic often makes it difficult and time-consuming to adjust job value and base salary to changed conditions. Flexible procedures have been designed to accommodate these needs.

Fourth, the larger the base pay proportion, the more organization members may perceive their pay as meaningful to satisfying important personal needs, according to the reflection theory on compensation. This theory holds that pay acquires meaning as it conveys information relevant to the self-concept of a person. Four meanings are distinguished: motivational properties, relative position, control, and spending. However, meaningfulness often diminishes as employees reach the high end of their salary scale without having much expectation of improving their level of pay. This implies that although base pay may motivate employees initially to stay in their jobs and to maintain at least an acceptable level of performance, this motivational force may wear out the longer employees remain in their jobs without noticeable changes in their level of pay.

4. PERFORMANCE-RELATED PAY

Research data and practical experience reveal that the introduction of a performance-related pay system is a time-consuming and rather complicated exercise. Evidently, performance-related pay may be successful in an organization, but there are also many instances where it fails to produce its intended effects. Many factors have to be considered, the most salient ones of which concern yardsticks of performance appraisal, anteceding conditions, size of pay differences, and some contingency variables (e.g., secrecy).

4.1. Performance Appraisal: Qualitative Versus Quantitative Yardsticks

In many companies, qualitative characteristics are used for the appraisal of individual performance such as

creativity, innovation, and customer orientation. These are often alternatively called behavioral yardsticks, personality statements, or competencies. A major problem is their abstract and often heterogeneous nature, without a direct relationship to observable work behaviors. They require a supervisor to "translate" daily and less frequent work events and actions into the framework of abstract person concepts without paying due attention to the interaction between the work situation and individual attributes. Thus, the use of these qualitative yardsticks is very much subject to bias. Because performance appraisal often occurs rather infrequently– (e.g., semiannually, annually), this method of appraising individual performance is vulnerable in practice. Because this performance appraisal result is tied to pay (either as a step on the salary scale or as a bonus), each department's supervisor is usually assigned a particular budget to control salary costs. The supervisor may award each employee with an equal share in the budget. However, performance-related pay is intended to reward differences in performance differentially. Characteristically, a normal distribution curve is applied to effectuate this, as Fig. 1 shows.

Figure 1 reveals that the great majority of employees earn an average bonus, whereas a few receive either a low bonus or a high bonus. Now assume that a supervisor wants to assign a relatively high bonus to some employees whose performance was beyond expectations. The restrictions posed by the budget necessitate that the supervisor assign a relatively low bonus to several other employees even though their performance often would not require this kind of intervention. How
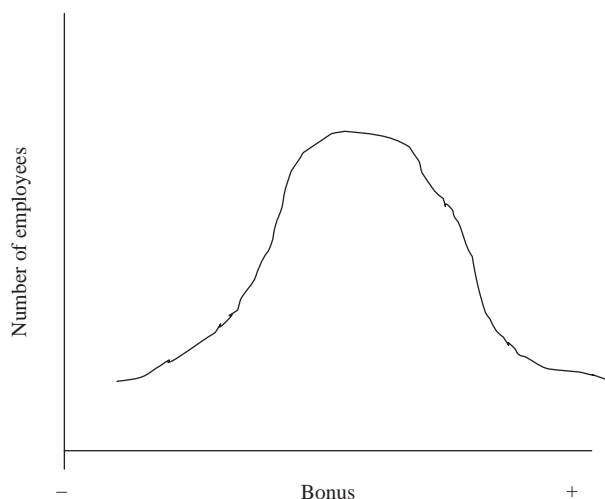

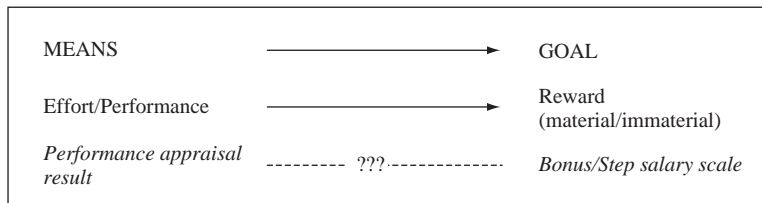FIGURE 1   Normal distribution curve.

FIGURE 2 Instrumentality perceptions according to expectancy theory and an absent "line of sight" as merit rating is applied.

should the supervisor get this message across? This example illustrates why research data on this system of merit rating (or competence pay) show time and again that employees hardly perceive a relationship between their performance appraisal result and their bonus. There is hardly any "line of sight," as Fig. 2 illustrates.

Figure 2 shows a core component of the expectancy theory on work motivation. Accordingly, the motivation to perform occurs as an individual perceives a relationship between his or her effort and performance level (expectancy) as well as a link between his or her performance level and one or more outcomes (instrumentality) that are attractive (valence). The figure highlights the principle of instrumentality and the often lacking line of sight—between the appraisal result and the resulting pay amount—as merit rating is applied. Not surprisingly, employees appear to be dissatisfied with their pay and consider the system (as well as the results it produces) to be unfair. The bonuses hardly recognize employees' performance contributions, and supervisors often wonder whether the bonuses can ever motivate employees to perform better.

There is no recipe for success. But when quantitative yardsticks are used, performance appraisal is a much less vulnerable process. Such yardsticks relate to outcomes of performance or to the pursuit of particular behavioral actions (e.g., prescribed behavior patterns in risky work situations). Performance outcomes can be set through defining a job's critical performance factors—two or three characteristics that represent the essence of a job (or job family) such as the number of billable training hours (for professional trainers) or the rate of recurring errors (for equipment repairers). Defining these factors is time-consuming; however, they can also be used for education and training, personal development, and career planning and can also be the focus of leadership behavior. Because critical performance factors should be vital to a job and sensitive to the job incumbents' decisions and actions, they lend themselves to performance pay as well. Research and practical experience show that their line of sight is usually adequate.

4.2. Anteceding Conditions

A performance-related pay system requires a careful phase of preparation. At least the following four conditions should be met.

4.2.1. Study of Work

Through analyzing work processes within a unit or department, it should become clear whether work activities are adequately tuned to each other, work methods are well chosen, and the workforce is sufficiently trained. Within this context, critical performance factors may be set.

4.2.2. Norm Setting

For each performance factor, a specific target ought to be set. Research on goal-setting theory has revealed that difficult and specific goals (for individual tasks with single goals) lead to high performance as employees accept these targets and are provided with frequent feedback.

4.2.3. Performance–Pay Link

Performance-related pay may be awarded when a target has just been met, is clearly within reach, or has been surpassed. The performance–pay link may be proportional or either less or more than proportional. Expectancy theory would recommend rewarding achieved results as frequently as possible.

4.2.4. Control by Employees

Mutual trust appears to be one of the cornerstones for getting performance-related pay under way. Employees (and managers) concerned should be involved in implementing a new system—in its design, in its introduction, and/or in its daily administration. The more employees believe that they are in control of a system, the more they tend to evaluate the system as fair.

4.3. Size of Pay Differences

Suppose that the salary an employee earned this month is several euros or dollars less than the amount he made last month. Chances are great that the employee will notify the salary administration department. However, if the employee's most recent paycheck would have shown a slightly higher salary than before, he probably would not have inquired about the difference. Upward and downward pay differences are not each other's reverse in their behavioral effects. Some studies show that a pay increase of less than 3% is usually not perceived as significant by individual employees (regardless of the level of their previous earnings). A pay increases of 5% or more is considered to be worthwhile "to go for," for instance, through better performance. However, if performance-related pay is applied, the size of the bonus appears not to be as important as the performance–bonus link per se.

Yet people at work often tend to compare themselves with others regarding their contributions to the work and the rewards they receive in return. Social comparison theory and equity theory have revealed that persons tend to compare their contributions–reward ratios with those of others who are in the most comparable situations (i.e., a referent). An employee who is comparing a valuable, relatively scarce attribute (e.g., his or her pay) with a referent wants to come out just a little bit higher. Such an outcome is quite motivating. But it hurts when the outcome is that the employee comes out a little bit lower. It is small differences that count.

4.4. Some Contingencies

Compensation systems and strategies operate within the social and cultural climate of an organization (and within the society at large). Consequently, many factors have been shown to affect compensation, one of which is pay secrecy. Many employers are quite open about their pay practices, whereas many others prefer to keep these secret. Whatever the arguments for such secrecy, it has an impact on the climate of trust and does not prevent organization members from making earnings comparisons among themselves. Research data show that such comparisons are unavoidably based on incomplete information, which may result in the misperception that the difference with the earnings of immediate subordinates is too small and that the difference with the earnings of superiors is too large. This may cause pay dissatisfaction. However, it is important to differentiate among three levels of "progressively open" information: the compensation system, its salary scales, and the level of individual earnings.

Also, differences in personality variables may be important. For example, individuals with higher scores on self-efficacy are more in favor of performance-related pay. Also, the higher the negative affect (i.e., the tendency to focus on negative life experiences that causes feelings of guilt and shame), the more disadvantageous effects a performance-related pay system will have.

5. CONCLUSION

Compensation in organizations provides a good example of multidisciplinary concerns. Compared with sociological, business administrative, and (especially) economic contributions, more psychological expertise would be most welcome. Recent work on the relationship between an organization's general business strategy and its compensation strategy calls for better theories (and more research) on the choice of particular pay patterns. Psychology may profit greatly from the groundwork laid in other disciplines.

See Also the Following Articles

Job Analysis, Design, and Evaluation

Further Reading

Dickinson, A. M., & Gillette, K. L. (1993). A comparison of the effects of two individual monetary incentive systems on productivity: Piece rate versus base rate pay plus incentives. *Journal of Organizational Behavior Management*, *14*, 3–83.

Gomez-Mejia, L. R., & Balkin, D. B. (1992). *Compensation, organizational strategy, and firm performance*. Cincinnati, OH: South-Western Publishing.

Lawler, E. E. (2000). *Rewarding excellence: Pay strategies for the new economy*. San Francisco: Jossey–Bass.

Lawler, E. E. (2003). *Treat people right*. San Francisco: Jossey–Bass.

Rynes, S. L., & Gerhart, B. (2000). *Compensation in organizations: Current research and practice*. San Francisco: Jossey–Bass.

Thierry, H. (2001). The reflection theory on compensation. In M. Erez, U. Kleinbeck, & H. Thierry (Eds.), *Work motivation in the context of a globalizing economy*. Mahwah, NJ: Lawrence Erlbaum.

Thierry, H. (2002). Enhancing performance through pay and reward systems. In S. Sonnentag (Ed.), *Psychological management of individual performance*. Chichester, UK: Wiley.

Competence at Work

Angelo S. DeNisi

Texas A&M University, College Station, Texas, USA

GLOSSARY

contextual performance Behaviors exhibited by employees that are not part of their formal job requirements but that contribute to overall organizational effectiveness; examples include volunteering for assignments, working with enthusiasm, and helping others at work.

frame of reference training A type of training in which raters are taught the definitions of various scale points on the rating scale rather than being taught to eliminate rating errors; the goal is to increase the accuracy of the ratings provided.

halo error A rating "error" manifesting itself as a high degree of consistency across ratings in various aspects of the job; this is assumed to occur due to a belief that good performance in one part of the job influences performance in other parts of the job, but eventually it became clear that this consistency might not be an error at all.

incident diary A log or record kept by the rater, noting all incidents of good and poor performance exhibited by an employee during a period of evaluation.

leniency error Another rating "error," in this case manifesting itself as ratings consistently above the rating scale mean value; eventually, it became clear that this too

might not be the result of a rating error but rather might reflect true levels of performance on the job.

Competence at work refers to the general evaluation of how individual employees (or teams of employees) contribute to the overall goals of the organization. Most typically, this evaluation is based on the judgments of some person (most often a direct supervisor), and formal evaluations are accompanied by documented ratings as well as feedback and discussion about ways in which to improve. Assessments of competence are often the basis for decisions about pay raises and promotions, but they can also be used as criterion measures against which to validate selection techniques. Underlying many of the purposes is the need to assess competence so as to help employees improve their performance on the job.

1. THE CONCEPT OF COMPETENCE AT WORK

An organization, whether private or public sector, hires employees to help the organization reach its goals. It is unusual, however, for every person hired to contribute to those goals equally, so it is important that we be able to assess the contribution of each employee (or group of employees) in the organization. There are a number of reasons why it is important that we are able to assess contributions in this way. One reason that is critical for employee selection is that we must be able to develop a

criterion measure against which to validate selection techniques. That is, for legal as well as practical reasons, it is important to know that the persons selected with a given technique will be those most likely to contribute to the organization's goals. But from a broader organizational perspective, we must be able to assess competence so that we can help employees to improve and expand the contributions they make to organizational goals. This is possible if we can assess relative strengths and weaknesses and then provide feedback about how to correct those weaknesses, and if we can demonstrate how improvements are tied to rewards such as pay increases, so that we can reward the employees accordingly. This requires that (a) we can define competence and the nature of that contribution and (b) we can assign some score or value to the contribution of each employee such that we can make some type of comparison across time and across employees.

The most traditional approach to defining competence at work is based on job analysis. That is, the purpose of job analysis is to specify what an employee is expected to do on the job, and the definition of competence on the job should probably be related to those requirements. But it is difficult to define competencies on the basis of specific tasks or operations carried out by employees. Such an approach would become extremely cumbersome given that we rarely have multiple employees on different jobs performing similar tasks. Thus, it makes more sense to define competencies in terms of more broadly defined contributions that underlie multiple tasks. As a result, in many cases, competencies are defined in terms of traits or characteristics of individuals that are believed to lead to success on a number of jobs. Furthermore, although a major portion of an employee's contribution will probably come in areas related to his or her job requirements, an employee can certainly make other types of contributions as well.

Any organization has certain tasks that must be carried out for the good of the organization such as employees being willing to help each other solve problems and employees volunteering to put in extra hours when crises arise. Any employee behaviors that make the workplace more pleasant are valuable. Yet none of these things is part of anyone's formal job description. Some scholars refer to these kinds of behaviors as organizational citizenship behaviors, whereas others simply refer to them as examples of contextual performance. Regardless of what they are called, they represent part of an employee's potential

contribution to the organization's goals, and they are part of any complete definition of competence. Although there has been disagreement over whether these things should be part of any evaluation of competence, most experts agree that these behaviors are considered when evaluation decisions are being made.

Finally, the operational definition of competence, usually expressed in terms of an evaluation instrument, can be stated in several different ways. The assumption that competence should be based on job analysis would lead to evaluation systems that focus on tasks that are to be carried out by the employees or on behaviors that employees should exhibit to be effective on the job. Others argue that the behaviors themselves are not as important as the outcomes of those behaviors. For example, we may be interested in whether salespeople follow up on sales calls, but what we really care about is whether or not they make the sales. Therefore, some have suggested that we define competency and assess contributions in terms of observable outcomes such as units produced and sales volume.

This approach can work only when the outcomes involved are under the direct control of the employees, and systems based on outcomes tend to work better when they are tied to specific goals that the employees must accomplish. That is, the exact nature of the contribution should be specified and should serve as a goal the employees try to reach during some time period. We could, for example, set a sales goal of 10,000 units or of $1 million. This would then define what we mean as competence at work; it would mean meeting that goal.

This approach has the benefit of defining expectations in the clearest of terms, and it allows employees to judge how well they are doing. It also makes the entire process somewhat less subjective (although the nature and level of goals are typically negotiable), and this should make it more acceptable to all parties involved. But this approach is not without its problems. It is critical that the organization state goals and objectives in the right terms and not just in terms that can be easily quantified. Otherwise, the organization runs the risk of having all of its employees meet their goals while the organization itself fails to meet its goals.

The final way in which competency has been defined at work is based on the presumed underlying traits and abilities needed to carry out the job. For example, regardless of the exact job duties, we might desire employees to demonstrate initiative on the job. Furthermore, we might believe that initiative will help those employees to carry out their assigned tasks and meet any meaningful goals. Defining competency and

assessing performance in this way allows the organization to use one measurement instrument across a wide variety of jobs, and this is the most commonly found approach to assessing performance. Although this approach leads to the most subjective of evaluation systems with standards that are difficult to define, there is reason to believe that there are certain underlying traits (e.g., dependability) that truly do differentiate outstanding employees from merely competent employees.

Each of these approaches to defining competence at work has its advantages and disadvantages. Although no one approach is clearly the best, the definition used will drive many of the management initiatives undertaken by the organization, so it is a critical decision. Thus, the key is to understand the implications of each approach and then decide on which one is best for the particular organization.

2. MEASUREMENT AND ASSESSMENT OF COMPETENCE

After the best working definition of competence is determined, the next major issue is determining how competence is to be measured. As noted previously, it is often not feasible to assess competence based directly on tasks performed. In addition, although there are instances where we can count units produced or monitor scrap rates, these instances are relatively rare, and as noted previously, the use of more objective measures can have the effect of focusing the employees' attention on the wrong things. As a result, most organizations use some type of rating scale for assessing competence, and many organizations spend a great deal of time experimenting with new forms.

Unfortunately, there is little evidence that any one type of rating instrument provides consistently better outcomes than other types of instruments. Therefore, it makes sense for an organization to use what works best for the organization and not to pay a lot of attention to what others are using. This issue is further complicated by the fact that there is no clear consensus as to what should be the criterion in any evaluation of measurement techniques. Furthermore, it is probably more important to focus on performance management techniques to try to improve competence than to focus on appraisal techniques designed to better assess competence.

Some of the efforts to develop better rating instruments have focused on the nature of the actual rating scales or on the number of points on those scales and the anchors used. Others have focused on exactly what is being assessed—whether behaviors, traits, or outcomes. But these efforts were hampered, in part, by the fact that there was no consensus as to how one should go about evaluating appraisal systems.

For a long time, the focus of improving appraisals was on the reduction of rating errors. These errors include things such as halo error (i.e., when someone who is rated highly on one aspect of a job is rated highly on all other aspects of the job) and leniency (i.e., when employee evaluations are inflated to be higher than they should be), and it was assumed that if there were less halo error and leniency in a set of evaluations, these would better reflect the true levels of competence of the employees being evaluated. But there is no reason to believe that a set of ratings that are lower, or are less intercorrelated, are necessarily better. It became clear that these "errors" might not be errors at all and that, in any case, they were generally unrelated to the accuracy of the evaluations provided.

Some scholars then began to assess rating accuracy directly. This made sense because it seemed reasonable to say that measures of competence were better if they accurately reflected the "true" levels of competence displayed by the employees. However, this approach requires a "true score" measure against which actual ratings can be compared, and such true scores are available only in laboratory settings.

More recently, it has become apparent that all of these efforts to reduce errors or improve accuracy are somewhat misplaced. In most organizations, the primary purpose of assessing competence is to allow employees to improve their levels of competence. In other words, the main purpose is to help employees improve the contributions they make to their organizations. Employees are more likely to improve if they perceive the assessment system as fair and accurate. To the extent that rating accuracy and the absence of what have been called rating errors can lead to perception of fairness and accuracy, these efforts are important. But they are important only as a means of getting to the ultimate goal—improvement.

Finally, there is the problem of how to deal with contextual performance when assessing competence. As noted previously, contextual performance refers to behaviors that benefit the organization but are not part of any employee's formal job requirements. The challenge here is not in deciding how to measure contextual performance (there are several scales available) but rather in deciding whether or not contextual performance should be assessed. If these behaviors

are included as part of a formal evaluation system, they will no longer be examples of contextual performance and instead will be part of the job duties. Would that be fair? Unfortunately, we do not have an answer yet.

3. BARRIERS TO EFFECTIVE ASSESSMENT

The solution to better assessment probably does not lie with the nature of the rating system used. Instead, the key component is the rater, that is, the person carrying out the evaluation of the employee's competence. Specifically, there are two issues involved. First, the rater must have the ability to provide fair and accurate evaluations. Second, and perhaps more important, the rater must be motivated to provide fair and accurate ratings. In other words, the rater must be able to recognize truly outstanding employees and then to make that recognition part of the public record.

Of course, the entire issue of what constitutes "fair and accurate ratings" is not a simple matter. We can define fair and accurate ratings as those that correctly describe employees' strengths and weaknesses and that assign the highest rating to the best performer, the next highest rating to the next best performer, and so on. There has been a great deal of research and discussion about how to determine rating accuracy, and many complicated formulas have been developed to help assess accuracy. However, the perception that a set of ratings is accurate is probably more important than the actual formula used to assess accuracy. It is this perception that will be critical for employees' belief as to whether a set of ratings is fair.

Attempts to improve rater ability have generally focused on training raters to be better judges. Consistent with the earlier view that reducing rating errors would lead to better evaluations, there was a great deal of interest in developing training programs to help raters reduce these errors. This interest gave way to developing training programs designed to help raters make more accurate ratings, usually referred to as "frame of reference" training. These programs emphasize helping raters to understand what constitutes good and poor performance in each area of competence being assessed. They are generally well received and have been found to improve rating accuracy in carefully controlled settings.

A different approach to improving rater ability deals with keeping performance logs or diaries. Formal appraisals are usually conducted no more than once or twice a year. Therefore, a rater must be able to recall the behaviors of each employee over a period of at least 6 months to provide accurate and meaningful evaluations. Clearly, if a rater "forgets" an incident, that incident cannot become part of the employee's evaluation. If the rater forgets enough of these critical behaviors, the evaluation will become less accurate. Furthermore, if there were some tendency to forget incidents involving good performance, the employee who received the evaluation would perceive it as unfair even though the rater was providing the fairest evaluation possible. A simple solution to this problem is to have the rater keep a diary or incident log over the evaluation period. The rater writes down (or somehow notes) every incident of good or poor performance and then uses this record to make an evaluation later. Although raters do not always keep these diaries as conscientiously as they should, there is evidence that such diaries are effective when they are used.

But even if the rater has the ability to provide fair and accurate evaluations, he or she might not always choose to act on that ability. Rater motivation is much more complex and less well understood than is rater ability. Why would a rater choose to be inaccurate? There are many reasons. Raters are human and so are prone to human biases. Many will favor people they like over people they dislike. Other forms of potential bias are along the lines of race, gender, and even age. Also, a rater may assign a lower rating than should be assigned just to establish his or her power to do so. Organizational policies and bureaucratic procedures may also make it easier for the rater to assign an inaccurate rating in some cases.

Raters will do what is in their own best interests. Assuming that they have the ability to provide fair and accurate ratings, they will actually do so only when they believe that either they will be rewarded for doing so or, at least, that they will not be punished for doing so. Organizational policies play a large role in this decision. For example, in some organizations, raters are required to provide excessive documentation if they give someone a truly outstanding evaluation. Faced with this requirement, raters may simply decide that it makes more sense to give a slightly lower evaluation.

In conclusion, raters are probably much more critical to the process of evaluating competence than is the instrument used. Raters must be trained, or otherwise helped, so that they have the ability to provide fair and accurate ratings, and policies must be such that raters will also be motivated to provide those ratings. If raters are capable of being fair and accurate as well as

motivated to be fair and accurate, the employees will be more likely to perceive the ratings as fair and accurate. Employees' reactions to these ratings, however, are critical to the effective management of performance at work.

4. REACTION TO ASSESSMENTS

As noted previously, the ultimate purpose of assessing competence at work is to motivate employees to improve their performance on the job and so increase their contributions to the organization. The issues of rater ability and motivation were discussed, and these factors are relevant for the employees as well. However, the question of employees' ability to improve performance relates more to selection and training activities in the organization and is not discussed in this article. The issue of employee motivation to improve, on the other hand, is central to any discussion of assessing and improving competence at work.

Following an evaluation, the desired outcome is for the employee to take the feedback he or she receives and to work to improve. If the feedback is positive (the employee is generally doing fine), this may require maintaining effort coupled with some refinement of where to allocate effort. This is typically not a problem. However, when the feedback is more negative (the employee needs real improvement), there are a number of problems that can occur.

A fairly typical reaction to negative feedback is simply to ignore the feedback. An employee might decide that the feedback is not fair or that the poor performance is somehow not the employee's fault. In any case, no improvement will follow this reaction. For the chances of acceptance to be high, the evaluation on which the feedback is based must be perceived as fair. That is, the evaluation itself must be seen as fair, and the procedures used by the rater to develop that evaluation must be seen as fair. In addition, the feedback must be delivered in a constructive way that does not threaten the self-esteem of the employee. Finally, the feedback should focus on the behaviors at work rather than on the person. Research suggests that feedback is not nearly as effective in changing behavior as is typically assumed, so it is critical that the employee accepts the feedback and works to improve.

During recent years, there has been a great deal of interest in a particular approach to competence assessment that is concerned with the issue of feedback. Multisource assessments, more commonly referred to as "360 degree appraisals," were designed to provide employees with feedback from peers, subordinates, and superiors and to compare this feedback with their own self-assessments. The assumption underlying this approach is that various individuals who have different relationships with an employee will focus on different behaviors. By providing feedback from all of the various sources, the employee should have a clearer picture of his or her strengths and weaknesses. There is actually little research about the effectiveness of 360 degree appraisals, and the results of the research that does exist are somewhat mixed. Nonetheless, the notion that individuals from various perspectives might have different views of an employee's competence, and that all of these may be useful, is an interesting idea.

5. DEVELOPING AND IMPROVING COMPETENCE

The penultimate goal of assessing competence is to improve competence. In truth, the ultimate goal of the process is to improve the performance of the entire organization. This must begin, however, with helping employees to accept and act on feedback. Once the employee is motivated to improve, the organization can take steps to aid in that improvement. In fact, if the organization does not take some steps in this direction, it is unlikely that real change will occur.

For example, if it is important to build and develop competence, feedback needs to be a regular thing. That is, feedback about performance should not occur just once or twice a year; it should occur constantly. Whenever an employee exhibits desirable or undesirable behavior on the job, someone should note this; in the case of undesirable behavior, the employee should also be told how to improve. Having supervisors set improvement goals with employees and regularly monitor progress toward those goals is also essential for real improvement. It is also important that supervisors and managers be constantly aware of cases where they might actually be punishing employees who perform well. A classic example is asking an employee to perform a "rush job" and then "rewarding" that employee by assigning him or her an additional job to perform.

It is hoped that if all individual employees improve their level of competence on the job, the organization as a whole will somehow be more effective. The truth is that this upward aggregation of improvement is not inevitable and requires close monitoring. Management, at the highest levels, must ensure that the goals and

directions set for individual employees at every level are consistent with, and supportive of, the organization's strategic goals. Therefore, it is critical that organizational goals be set with some recognition of the human resources available to meet them and that tactics for improving competence consider how improvement can help to serve those strategic goals. It is also important for high-level managers to recognize that they must be able to leverage the competence they have. That is, if an organization has extremely creative employees, its managers should pursue strategies that emphasize the importance of creativity. In this way, an organization can ensure that improvements in individual competence can translate to improvements in organizational effectiveness.

6. CONCLUSION

An organization improves its effectiveness by improving the competence of individual employees at work so that these individuals can make greater contributions to the organization's strategic goals. This process begins with the definition of competence at work and what exactly is included in that definition. Once a definition is selected, the organization must develop the means by which to determine the current level of employee competence versus where the employee needs to be. Although this might sound simple, there are a number of barriers to effective assessment. Some of them stem from the characteristics of the persons doing the assessment, whereas others stem from organizational policies that may (unintentionally) make this assessment more difficult.

But it is not enough to develop sound assessment techniques. Those assessments must be fed back to the employee in ways that will encourage him or her to work toward improvement. Organizational policies related to goal setting and regular feedback can help to ensure that this occurs. In addition, it is important to coordinate goals and incentives at every level in the organization so that improved competence at the individual level can translate to improved organizational effectiveness.

See Also the Following Articles

Job Analysis, Design, and Evaluation ■ Organizational Participation

Further Reading

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include assessments of contextual performance. In N. Schmitt, & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). San Francisco: Jossey–Bass.

Colquitt, J., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86, 425–455.

DeNisi, A. S. (2000). Performance appraisal and control systems: A multilevel approach. In K. Klein, & S. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 121–156). San Francisco: Jossey–Bass.

Kluger, A. N., & DeNisi, A. S. (1996). The effects of feedback interventions on performance: Historical review, meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87, 72–107.

Murphy, K. R. (1991). Criterion issues in performance appraisal research: Behavioral accuracy versus classification accuracy. *Organizational Behavior and Human Decision Processes*, 50, 45–50.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Thousand Oaks, CA: Sage.

Competition in Sport

Daniel Gould

Michigan State University, East Lansing, Michigan, USA

Cristina Rolo

University of North Carolina, Greensboro, North Carolina, USA

GLOSSARY

burnout A psychological, physiological, and/or emotional withdrawal from sport due to exhaustion, depersonalization, and low accomplishment feelings resulting from chronic stress.

competition According to Martens, a process whereby an individual's performance is compared with some standard of excellence in the presence of at least one other person who is aware of the criterion for comparison.

state anxiety Feelings of apprehension and tension associated with activation of the organism.

stress An athlete's perception of the imbalance between the environmental demands placed on him or her and the athlete's response capacity and resources for meeting those demands.

trait anxiety A personality disposition that predisposes individuals to perceive evaluative contexts as more or less threatening and to respond with varying levels of state anxiety.

Competition is a process that has important positive and negative consequences for athletes. Positive consequences can include increased confidence, motivation,

and satisfaction, whereas negative consequences can include stress and burnout. The degree to which competition has positive or negative consequences depends on the competitive context (e.g., the importance placed on performance and uncertainty) and the personality (e.g., trait anxiety, self-esteem) and perception of the athlete involved. Positive coaching strategies and stress management techniques can be used to reduce or cope with competitive stress.

1. INTRODUCTION

1.1. Competitive Sport: A Worldwide Phenomenon

Competitive sport is a worldwide phenomenon of enormous popularity. Epitomized by the Olympic Games and World Cup soccer championships, millions of individuals watch these events with tremendous enthusiasm. Moreover, countless others participate in various levels of competitive sport every year.

Competitive sport is not without controversy. Some argue that it is an excellent training ground for psychological attributes such as leadership, confidence, and teamwork, whereas others suggest that it leads to excessive anxiety, immoral behavior, and aggression.

Because of the persuasiveness of competitive sport, better understanding competition and the competitive process has been an important task for those in applied psychology. Three decades of sport psychological

research on competition have provided a solid under-standing of the area.

1.2. Purposes

This article has three purposes. First, competition is defined and the competitive process is discussed. Second, the issue of whether sport competition has positive or negative effects on participants is examined. Finally, competitive stress, its sources, and its consequences are discussed.

2. COMPETITION AND THE COMPETITION PROCESS

Based on the groundbreaking work of Deutsch in 1949, early definitions of competition focused on situations where rewards were distributed unequally based on the performance of participants. Although this definition did much to guide early research in the area, its shortcomings (e.g., the fact that individuals do not always define rewards in the same way and, hence, that a person might lose an event but define his or her personal improvement as rewarding) led to the development of other definitions.

Today, competition is defined as a social process that encompasses a sequence of stages rather than a single event. Martens did much to advance the scientific study of competition when, in 1975, he defined it as a process where "an individual's performance is compared with some standard of excellence in the presence of at least one other person who is aware of the criterion for comparison." This differs from the reward definition in that the person is the focal point of the competition process and can influence the relationship among the various stages of competition. Primary emphasis is also placed on the individual's perception of the competitive context and on how personality factors influence one's competitive experience.

Standards of excellence involved in the competition process can be self-referenced (e.g., set a personal best in a 100-meter swim competition) or other-referenced (e.g., beat a particular individual in the competition). Considerable attention in contemporary sport psychology focuses on the utility of making self- versus outcome-oriented comparisons. Self-referenced standards are often emphasized because they are more in an individual's control, and have been found to be associated with long-term motivation and greater satisfaction, compared with other-referenced comparisons.

3. A MODEL FOR UNDERSTANDING COMPETITION

Guided by the process definition of competition, Martens also developed a four-stage model of competition (Fig. 1). An inspection of this model reveals that the objective competitive situation (Stage 1) begins the competitive process and consists of a situation where an individual's performance is compared with others in the presence of least one other individual. For example, for most individuals, entering a gymnasium and seeing a basketball at midcourt and one's name on the scoreboard would signal a competitive situation. However, not everyone would perceive this situation as competitive, for example, an elderly person who has no interest in or knowledge of basketball. Hence, Stage 2 of the process involves an individual's subjective perception of the objective competitive situation. Research has shown that an individual's personality orientations, particularly his or her levels of competitiveness or achievement goals (i.e., orientation to judge ability relative to oneself or a comparison with others), influence the degree to which the person sees situations as competitive.

After an individual appraises a situation as competitive, he or she may approach or avoid it or may become motivated, anxious, or excited. Thus, Stage 3 of the model focuses on an individual's psychological–physiological response (e.g., anxiety, increased muscle tension, enhanced motivation). Interestingly, research also shows that it is not the absolute level of anxiety

Stage 1: Objective Competitive Situation
(situations where one's performance is compared with others)

↓

Stage 2: Subjective Competitive Situation
(the degree to which one perceives objective
competition situations as evaluative)

↓

Stage 3: Response
(physical and psychological responses to competition such
as increased state anxiety and motivation)

↓

Stage 4: Consequences
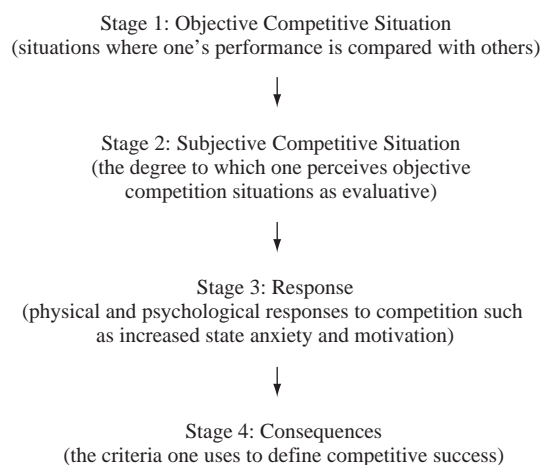(the criteria one uses to define competitive success)

FIGURE 1 Martens' model of competition. Reprinted from Martens (1975).

that influences an athlete's performance; rather, it is whether the individual views that increased anxiety as facilitative and helpful or as debilitative and a hindrance to performance. The athlete's perception, then, not only influences whether he or she views competition as anxiety provoking but also influences whether that anxiety is seen as helpful or hurtful.

Finally, the positive or negative consequences are the fourth stage in the competition process. An individual wins or loses on the scoreboard or perceives that he or she performed well or not. However, as with the three preceding stages, an athlete's perception of the consequences is more important than the objective outcome.

This model of competition is important because it shows that competition can best be understood as a series of stages. Moreover, it places an athlete at the center of the competitive process, with his or her personality dispositions interacting with environmental considerations.

4. POSITIVE AND NEGATIVE EFFECTS OF COMPETITION

The competitive ethic is a driving force in contemporary sport. It is common to hear coaches, athletes, and sports journalists say positive things such as "competition brings out the best in people" and "she is a fiery competitor." At the same time, an overly competitive, win-at-all-costs mentality is blamed for violence, rules infractions, and unsportsmanlike behavior. Hence, sports competition appears to be a double-edged sword, cutting both ways and having both positive and negative effects on participants.

Most researchers do not view competition as inherently good or bad today. Research shows that it can have both positive and negative effects on participants. For instance, in some studies, competition has been shown to lead to negative effects such as increased aggression and decreased performance. In other investigations, competition has been found to facilitate motivation and lead to improved performance. Whether these effects are positive or negative depend greatly on the competitive context and the emphasis that sport leaders and coaches place on competition and its meaning. For example, the quality of adult leadership has been shown to be a crucial determinant of whether competition has positive or negative outcomes for children. In their classic youth sports coaching research, Smoll and Smith found that children playing for coaches trained in a positive approach (i.e., who focused on encouragement and giving praise) exhibited

higher self-esteem and lower anxiety and were less likely to drop out of sport. When coaches were not trained to be positive and encouraging, children involved in competitive sport did not experience increased motivation, lower anxiety and enhanced self-esteem.

Much of the sport psychological research focusing on competition has been conducted with young athletes in entry-level programs. Less research on competition has been conducted at different levels of competition (local, national, and international). However, recent research with elite athletes shows that these outstanding performers are highly competitive. Their standards of comparison are both self- and other-referenced, they are both intrinsically and extrinsically motivated, and they have high perceptions of ability.

Champion athletes have also been found to approach competition differently across various phases of their careers. For instance, in the classic research by Bloom in 1985, elite athlete development was retrospectively traced across the careers of champion athletes. Bloom's research, as well as more recent research, found that these individuals did not begin their careers with an emphasis on competitive outcome (i.e., winning medals and defeating others). Rather, the focus was on having fun, learning fundamentals, and being active. It was only later that they approached competition in a very serious fashion. In addition, for much of their careers, they focused on long-term development rather than short-term competitive results.

5. COMPETITIVE STRESS IN SPORT

One of the most studied aspects of competitive sport has been its association with competitive stress. Levels of stress experienced by young athletes, sources of stress, and managing stress all have been topics of considerable interest to sport psychology researchers.

5.1. Stress in Sport

Stress has been defined as an athlete's perception of the imbalance between the environmental demands placed on him or her and the athlete's response capacity and resources for meeting those demands. For example, a golfer may face a situation where he or she needs to sink a crucial 15-foot putt to win a tournament. If the golfer perceives the demands as exceeding his or her capabilities, the result is increased competitive state anxiety, that is, feelings of apprehension and tension accompanied by physiological activation. Furthermore,

heightened levels of state anxiety, especially if they are perceived as debilitative, have been associated with poor performance and lower levels of enjoyment and satisfaction.

5.2. Sources of Stress in Athletes

Identifying sources of stress experienced by athletes engaging in competition has been a topic of interest to sport psychologists. This research has generally shown that although there are a variety of specific stress sources that an athlete can experience (e.g., self-doubts about performance capabilities, team selection), these fall into two general situational categories: (a) the importance placed on performance and (b) uncertainty. Specifically, the more importance that an athlete perceives is placed on an event, the more state anxiety is experienced. Similarly, the greater the degree of uncertainty (whether about performance or nonperformance issues) that the athlete perceives, the more competitive state anxiety is experienced.

However, situational factors are not the only class of factors influencing competitive stress responses. Consistent with Martens' model of competition, personality dispositions have been associated with elevated stress responses. Specifically, an athlete's level of trait anxiety (i.e., a personality orientation that predisposes the individual to view social evaluation and competitive contexts as threatening) influences his or her level of state anxiety. Performers with high trait anxiety consistently respond with greater levels of state anxiety in competitive situations. Self-esteem has also been consistently associated with levels of competitive state anxiety experienced in socially evaluative sport contexts. In competitive situations, athletes with low self-esteem experience higher levels of state anxiety than do those with high self-esteem. Finally, more recent studies have found relationships between increased anxiety and high levels of hardiness, perfectionism, and self-presentation concerns as well as low social support, although these findings need to be further verified in additional studies.

5.3. Are Young Athletes Placed under Too Much Stress?

For a number of years, researchers have been concerned with the levels of stress experienced by athletes, especially young athletes engaged in competitive sports. Thus, researchers have compared the amounts of state anxiety experienced during practices with those experienced during competitive sport situations. The thinking behind these comparisons is that competition is more anxiety provoking than are practices, but it is important to note that at times practices can include considerable social evaluation (e.g., team selections). Although the results of these studies have consistently revealed that most young athletes experience more anxiety during competitions than during practices, the most important finding is that the vast majority of children do not experience excessive levels of state anxiety during competitions. However, certain children in certain situations do experience high levels of competitive stress and anxiety.

For example, in a set of classic studies of competitive youth sport participants, Scanlan and colleagues assessed the state anxiety levels of youth sport participants during practices (i.e., nonevaluative contexts) and compared these levels with those during competitions (i.e., evaluative contexts). A variety of personality and background factors were also assessed. Results of this research led to the general conclusion that most young athletes did not experience excessively high levels of state anxiety during competitions. Certain children in certain situations did experience high levels of state anxiety. These children were characterized by high competitive trait anxiety, low self-esteem, less fun, less satisfaction with performance, low personal performance expectancies, and worries about failure and adult evaluation.

In a related area of research, "burnout" or withdrawal from competitive sport has been studied. In these studies, burnout has been defined as a psychological, physiological, and/or emotional withdrawal from sport due to exhaustion, depersonalization, and low-accomplishment feelings resulting from chronic stress. Results reveal that a small percentage of sport participants experience burnout of sport and that chronic stress plays an important role in the burnout process. Moreover, the chronic stress results from physical overtraining and/or psychological factors such as overbearing parents or coaches, a lack of athlete multidimensional identity development, a perfectionistic personality orientation, high trait anxiety, and low self-esteem.

5.4. Managing Competitive Stress

Because competitive sport can be stressful, and high levels of stress have been associated with inferior performance and decreased enjoyment and satisfaction, researchers have been very interested in helping athletes to manage stress.

One important but often overlooked class of techniques for managing competitive stress is environmental engineering. With environmental engineering

techniques, coaches, adult leaders, and significant others can influence athlete stress levels by increasing the importance placed on competition (e.g., giving a fiery pregame pep talk, repeatedly emphasizing the importance of victory) or by creating uncertainty (e.g., not announcing starting lineups for games, basing liking for a child on performance). Or, they can reduce competitive stress by reducing the importance placed on performance (e.g., not giving a pregame pep talk, not repeatedly emphasizing the importance of victory) or by increasing certainty (e.g., announcing starting lineups early, not basing liking for a child on performance). Training youth sport coaches to adopt a positive and encouraging coaching orientation, rather than a negative or critical one, can reduce the levels of competitive stress experienced.

A number of techniques, mirroring general stress management research in psychology, have been used successfully by athletes to manage their competitive stress. These include cognitive anxiety reduction strategies (e.g., imagery, meditation), somatic anxiety reduction techniques (e.g., progressive muscle relaxation, breath control, biofeedback), and multimodal techniques (e.g., stress inoculation, cognitive–affective stress management training). Competitive athletes have also been found to use a variety of problem-focused (e.g., time management, goal setting) or emotion-focused (e.g., controlled breathing, relaxation training) coping techniques to control the stress of competition.

Relative to the utility of teaching stress management to athletes, sport psychologists emphasize the need for long-term training/practice and the incorporation of these techniques into the actual competitive setting (e.g., learning how to relax while an athlete actually runs vs using relaxation techniques the night before a race). Finally, research has also revealed that there is an optimal recipe of emotions that lead to superior sport performance, so that using stress management techniques to eliminate all stress is counterproductive. Thus, competitive athletes must know their optimal levels of emotional arousal and arousal-related emotions needed for best performance and then use stress reduction and enhancement strategies accordingly.

See Also the Following Articles

Anxiety and Optimal Athletic Performance ■ Decision Making in Sport ■ Goal Setting and Achievement Motivation in Sport ■ Overtraining and Burnout in Sports ■ Performance Slumps in Sport: Prevention and Coping ■ Psychological Skills Training in Sport ■ Self-Confidence in Athletes ■ Sport Psychology, Overview ■ Stress

Further Reading

Bloom, B. S. (Ed.). (1985). *Developing talent in young people.* New York: Ballantine.

Deutsch, M. (1949). An experimental study of the effects of cooperation and competition upon group processes. *Human Relations*, 2, 199–231.

Gill, D. (2000). *Psychological dynamics of sport and exercise* (2nd ed.). Champaign, IL: Human Kinetics.

Gould, D. (1993). Intensive sport participation and the prepubescent athlete: Competitive stress and burnout. In B. R. Cahill, & A. J. Pearl (Eds.), *Intensive sport participation in children's sports* (pp. 19–38). Champaign, IL: Human Kinetics.

Hanin, Y. L. (Ed.). (2000). *Emotions in sport.* Champaign, IL: Human Kinetics.

Horn, T. (2002). *Advances in sport psychology* (2nd ed.). Champaign, IL: Human Kinetics.

Jones, G., Hanton, S., & Swain, A. (1994). Intensity and interpretation of anxiety symptoms in elite and non-elite sports performers. *Personality and Individual Differences*, 17, 657–663.

Martens, R. (1975). *Social psychology of sport.* New York: Harper & Row.

Scanlan, T. K. (1986). Competitive stress in children. In M. R. Weiss, & D. Gould (Eds.), *Sport for children and youth* (pp. 113–118). Champaign, IL: Human Kinetics.

Smith, R. E., Smoll, F. L., & Curtis, B. (1979). Coach effectiveness training: A cognitive–behavioral approach to enhancing relationship skills in youth sport coaches. *Journal of Sport Psychology*, 1, 59–75.

Smoll, F. L., & Smith, R. E. (Eds.). (2002). *Children and youth in sport* (2nd ed.). Dubuque, IA: Kendall/Hunt.

Weinberg, R. S., & Gould, D. (2003). *Foundations of sport and exercise psychology* (3rd ed.). Champaign, IL: Human Kinetics.

Conflict within Organizations

Michael Ryan and Stephanie L. Castro

Florida Atlantic University, Boca Raton, Florida, USA

GLOSSARY

conflict episode A single engagement between two parties in conflict over some issue.

conflict process The series of activities from initiating conflict to resolving it.

distributive outcome Zero-sum resolution to a conflict in which the outcome is perceived as fixed and the purpose of the conflict is to distribute the outcome in some proportion between the parties (e.g., 50/50, 60/40, 100/0).

external conflict Conflict existing between an individual or group and another individual or group.

goal conflict Conflict between parties where one party has a goal that is mutually exclusive with a goal of another party.

integrative outcome Resolution to a conflict in which the parties work to create an outcome that is compatible with both parties' goals.

internal conflict Conflict existing within an individual who is experiencing mutually exclusive desires or beliefs.

interpersonal conflict A state that exists between two individuals with valued interests in which one individual's (or each individual's) interests are violated, or in danger of being violated, by the other individual.

judgment conflict Conflict between parties where one party's assertion about something is inconsistent with another party's assertion about it.

normative conflict Conflict between parties where one party perceives that the other party has exhibited behavior that is believed to be wrong.

Organizational conflict is a state that exists between two parties within an organization or between organizations, each with valued interests, in which one party's (or each party's) interests are violated, or in danger of being violated, by the other party.

1. INTRODUCTION

Both within and between organizations, individuals interact regularly and almost as regularly experience conflict with one another. Sometimes the conflict between individuals is minor and unnoticed, whereas other times it is heated and drawn out. Whatever the case, the characteristics of conflict between individuals are important to understand for the sake of the individuals and the organizations where they work. This article provides a general understanding of interpersonal conflict in organizations and highlights some of the most recognized ideas that have been presented in research on the subject.

447

Interpersonal conflict is a state that exists between two individuals with valued interests in which one individual's (or each individual's) interests are violated, or in danger of being violated, by the other individual. It is an external form of conflict, as opposed to internal conflict where two or more of an individual's interests are mutually exclusive (e.g., role conflict, conflict of interest).

Interpersonal conflict occurs between parties with opposing interests and can take three forms: normative, judgment, and goal. Normative conflict exists when one individual is offended by another individual's behavior and believes that the latter should have behaved differently. Judgment conflict occurs when the individuals disagree over a factual issue (e.g., how a particular event transpired). Goal conflict is a situation in which two individuals are seeking mutually exclusive outcomes (e.g., how a slice of cake will be divided between them).

In 1992, Thomas, an extremely influential figure in conflict research, divided conflict into four factors that affect one another directly or indirectly: structural conditions, the conflict process, outcomes, and third-party intervention. Structural conditions make up the context of the situation in which conflict occurs and include the relatively stable factors of the parties and their surroundings. When conflict is triggered, the conflict process ensues with interactions between the parties involved. Both experiences and behaviors of the parties affect the process. Conflict outcomes are the result of the conflict process (specific episodes or series of episodes) and can include task outcomes (relevant to the matter of the conflict) and social system outcomes (relevant to the organization, the relationship between the parties, etc.). During the process, a third party may intervene. Third-party intervention is the involvement of an individual or a group whose primary concern is the resolution of the conflict as opposed to the substantive concerns of the conflicting parties.

Some third parties may, on resolution, change the structural conditions of the context, thereby influencing the resolution of future conflict. An example might be a manager who helps to resolve a dispute between two employees and then creates a policy to handle similar disputes in the future. Aside from these primary effects, feedback effects are likely to occur. For instance, the social system outcomes of a conflict episode are likely to alter structural conditions such that no two conflict episodes occur in identical environments.

2. THE CONFLICT PROCESS

When two individuals are experiencing conflict, the conflict process is the most apparent and explicit aspect of all that goes on between the parties. They are likely unaware of the contextual factors and their individual characteristics that set the conflict in motion. Although they have an idea of what they want out of the conflict, they are probably not considering all of the effects that the conflict will have beyond their individual goals. It is likely that each party is most interested in how to behave, and in how the other party should behave, in dealing with the conflict. The series of activities that run from initiating conflict to resolving it is referred to as the conflict process. The most popular research on the conflict process suggests that the conflict process is best observed at the level of the conflict episode. A conflict episode is a single engagement between two parties in conflict over some issue.

Two researchers developed and presented influential models of the conflict process: Pondy in 1967 and Thomas in 1992. Although both models have similar aspects, Thomas's model gives attention to additional characteristics of the process such as the interaction of emotions, rational reasoning, and normative reasoning. The main aspects of this model are awareness that conflict exists, a series of thoughts and emotions leading to intentions that determine how the conflict will be handled, and interaction between the parties beginning with the first party's behavior and continuing with the other party's reaction. This leads to a continuous exchange cycle between the parties through thoughts/emotions, intentions, and behavior that ends with resolution and outcomes.

Thomas defined conflict as "the process that begins when one party perceives that the other has negatively affected, or is about to negatively affect, something that he or she cares about." At some point before an individual becomes involved in handling conflict with another person, that individual must first recognize that conflict exists. This point of first recognizing conflict is referred to as awareness. There will often be an event that triggers awareness such as an action or a statement made by the opposing party.

On becoming aware of conflict, an individual then begins to make sense of the situation by considering what is at stake and how the process should be handled. One aspect of this sense-making is defining the issue of the conflict. In addition to recognizing the issue at stake, an individual will begin to consider what

outcomes are likely to result in a resolution that is acceptable to both parties.

Having defined the issue and considered possible settlements, an individual will then begin to reason both rationally and normatively in considering ways in which to handle the conflict that (a) are likely to result in efficiently and effectively achieving a beneficial resolution and (b) are perceived as acceptable and fair by the social environment. For example, if two individuals are in conflict over purchasing the last of a particular toy in a store before Christmas, the issue is apparent: Which of the two will get to purchase the toy? Rationally, one party may consider that attacking the other party and stealing the toy would efficiently and effectively result in a beneficial outcome. However, normatively, the social environment would not approve of this behavior. Therefore, the party should consider a more acceptable strategy for acquiring the toy.

The writer Oscar Wilde once said, "Man is a reasonable animal who always loses his temper when he is called upon to act in accordance with the dictates of reason." Accordingly, individuals engaged in conflict are often affected by their emotions in addition to considering rational and normative reasoning. Both positive and negative emotions can affect how the opposing party's behavior is interpreted and can amplify or dampen the influence of rational and normative reasoning. All of this will filter down to influence behavior. For instance, anger can result in aggression and negative interpretations of behavior, anxiety can lead to withdrawal, and innocent humor can decrease the negative effects of aggression and anger.

The interaction of an individual's reasoning and emotions leads to a set of strategic and tactical intentions for handling the conflict episode. Strategic intentions refer to the overall goal one party has for the situation such as cooperating with or competing against the other party. Research on strategic intentions often refers to them as conflict management strategies and is discussed in the next section. Tactical intentions, on the other hand, are more specific goals or plans for achieving a beneficial outcome such as using a calm tone in presenting a case to keep the other party from becoming angry and focusing on the benefits that a proposed solution has for the other party. These intentions serve as guides for an individual's behavior in managing conflict.

The intentions that an individual forms for dealing with a conflict episode will influence that party's behavior. After the first party expresses behavior, the other party experiences a similar process of reasoning and experiencing emotions, forming intentions, and expressing behavior as a response to the interpretation of the first party's behavior. This is not necessarily a turn-based activity, although it can be in certain negotiation situations. Each party responds to the other party through the stages of thoughts/emotions, intentions, and behavior in a continuous interaction between the parties until some outcome is reached. This might not be the end of the conflict between these parties; however, it is the end of the conflict episode. The outcomes of conflict are discussed in a later section.

3. CONFLICT MANAGEMENT STRATEGIES

Five strategies for managing conflict have been used throughout numerous research studies on conflict: avoiding, accommodating, competing, compromising, and collaborating. These strategies are coordinated on a two-dimensional grid with the axes labeled either "concern for own interests" and "concern for other's interests" or "assertiveness" and "cooperativeness" (Fig. 1). Thus, avoiding would represent either a low concern for self-interests and for other's interests or low levels of assertiveness and cooperation, respectively.

The first conflict management strategy is accommodating, also referred to as obliging. Figure 1 shows that this strategy is low in assertiveness and high in cooperativeness—a lose–win strategy. In this case, one party chooses to allow the other party to satisfy its interests
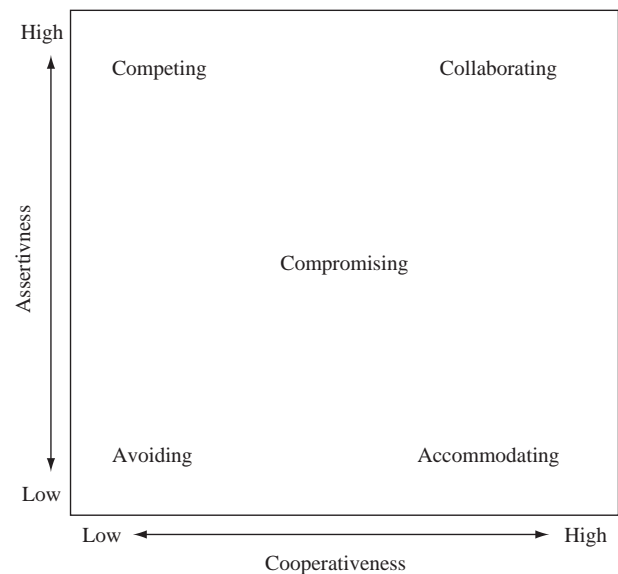

FIGURE 1   Conflict management strategies.

completely. One party may choose accommodation for several reasons. For example, the party may be unwilling to compete, may want to minimize losses, may consider the other party's interests as more valuable, or may seek to prevent damage to the relationship between the parties. Accommodating can be an active strategy in which one party helps to satisfy the other party's interests. It can also be a passive strategy in which inaction results in the satisfaction of the other party's interests.

The second conflict management strategy is competing or dominating. This is the opposing strategy to accommodating and is classified as high in assertiveness and low in cooperativeness—a win–lose strategy. Competing is an active strategy in which one party attempts to satisfy its own interests, usually by preventing the other party from satisfying its interests when both parties' interests are perceived as mutually exclusive. Competing may be more effective when quick and decisive action is vital, unpopular but important decisions must be made, an organization's welfare is at stake and the right course of action is known, and/or the other party takes advantage of noncompetitive behavior.

Compromising or sharing is the third strategy shown in Fig. 1. In this case, one party is moderately assertive and cooperative. In contrast to either party winning or losing, compromising seeks a "middle ground" in which each party wins in part and loses in part. Compromising may be perceived as a weaker form of collaborating, but the use of compromising assumes a fixed amount of resolution that must be distributed, whereas collaborating seeks to determine both the size of the resolution and how it is to be divided. Compromising may be effective when opponents with equal power are committed to mutually exclusive goals, the issue is complex and at least a temporary settlement must be reached, interests cannot be fully sacrificed but there is no time for integration, collaborating and competing fail, and/or goals are important enough to fight for but not so important that concessions cannot be made.

The fourth conflict management strategy is avoiding. This strategy reflects low assertiveness and low cooperation—a lose–lose situation. With this strategy, one party chooses not to engage the other party in resolving the conflict in such a way that neither party is able to satisfy its interests. Avoiding may be a good strategy when an issue is trivial, there is no chance to satisfy a party's own interests, the other party is enraged and irrational, and/or others can resolve the conflict more easily.

The last of the conflict management strategies is collaborating or integrating. As the term implies, this strategy has an integrative focus. Parties using the collaborating strategy are classified as being both highly assertive and cooperative—a win–win attitude. In this case, the two parties attempt to work with each other to develop a resolution that will completely satisfy the interests of both parties. Although much of the research on conflict has proposed that collaborating is always the ideal conflict management strategy, others have suggested that there may be situations in which other strategies are more effective. Collaborating may be a more effective strategy when each party's interests are too important to be compromised, full commitment to the resolution by both parties is desired, and/or there is sufficient time to resolve the conflict by integrating both parties' interests.

Research on conflict management strategies has comprised a large portion of the total research on interpersonal conflict. The five strategies outlined here have become somewhat canonical in conflict research. Much research has validated the existence of these strategies and shown the advantage of a two-dimensional model over a one-dimensional (two-strategy) model, but there have been no remarkable attempts to improve on or add to this model of conflict management. Although there is much room for more research to be done in this area of interpersonal conflict, it has not suffered from a lack of attention as other areas of conflict have.

4. STRUCTURAL CONDITIONS AFFECTING CONFLICT

The conflict process does not occur in a void; there are structural conditions that affect the process. Both the contextual situation and the parties involved have characteristics that may affect the various factors in the conflict process such as the choice of a conflict management strategy, inclusion of a third party, and norms for behavior. In 1992, Thomas suggested, "Whereas process refers to the temporal sequence of events that occurs in a system, structure refers to the parameters of the system—the more or less stable (or slow-changing) conditions or characteristics of the system that shape the course of those events." Research has revealed several structural factors in both individuals and organizations that may have an effect on the conflict process. For instance, Wall and Callister's review of conflict literature in 1995 discussed several individual and issue-related factors affecting conflict.

4.1. Individual Characteristics

Individuals constitute the primary parties in interpersonal conflict; therefore, it is likely that characteristics of those individuals will influence how the conflict episode is played out. It could be that differences in characteristics between individuals are the very reason for a conflict episode, for example, when two team members get into an argument because one is task oriented and the other is more relationship oriented. This difference can be the root of a series of events that result in conflict. Similarly, these differences may lead to each party handling the episode in a different way. Thus, individuals have stable or slow-changing characteristics that can influence conflict and the conflict process.

One characteristic that affects conflict is an individual's culture (i.e., societal norms). Conflict is seen by some cultures as beneficial but is seen as detrimental by others. Thus, more assertive strategies are likely to be preferred by individuals from cultures seeing conflict as beneficial. Other individual characteristics affecting conflict are goal characteristics, stress, and the need for autonomy. An individual with high aspirations or extremely rigid goals is more likely to engage in conflict due to increased awareness of possible impediments to his or her goals. In addition, an individual whose goals are interdependent with another person's goals is more likely to engage in conflict because the latter person's behavior has an impact on the former person's ability to achieve a goal. Stress may also lead to greater conflict due to the tension that an individual experiences, much like a rubber band pulled tight and ready to snap. Finally, if an individual has a high need for autonomy and another party makes demands that infringe on that autonomy, there is an increased likelihood of conflict.

Another individual characteristic that affects conflict is personality. For example, "Type A" and "Type B" personalities are likely to manage conflict in different ways given that Type B individuals are less competitive, less temperamental, and more patient. Studies by Antonioni in 1998 and Moberg in 2001 examined the effects of the "Big Five" personality factors—Extraversion, Openness, Agreeableness, Conscientiousness, and Neuroticism—on preference of conflict management strategy. For instance, avoiding seems to be preferred by individuals with low Extraversion and low Conscientiousness, collaborating is preferred by those with high Extraversion and high Conscientiousness, and competing is associated with low Agreeableness.

Self-esteem has received considerable attention in conflict research. Research has shown that individuals with high self-esteem value satisfying both their own concerns and their counterparts' concerns. These individuals are also more confident that collaboration is possible and are less likely to become defensive, a condition that negatively affects collaboration. Thus, individuals with high self-esteem are more likely to use collaboration in managing conflict. Other individual factors affecting conflict include locus of control (internal may provide greater confidence for resolving conflict), dogmatism (high levels may limit the flexibility of resolution possibilities), and stage of moral development (higher stage individuals are more likely to use collaboration). The individual characteristics presented here do not represent a comprehensive list of individuals' structural variables that may affect conflict. However, these are some that have been addressed in research. Primarily, the discussion presented here suggests that it is important to consider more than the issue when trying to understand a conflict episode. The individuals themselves may have characteristics that contribute to both the triggering of the conflict episode and its management.

4.2. Organizational Characteristics

Conflict does not occur in a vacuum; rather, it occurs in some sort of interactive setting such as the workplace. Much of the research presented here (and in general) has concentrated on conflict occurring in an organizational context. Thus, much research has focused on examining the characteristics of organizations that influence conflict and its management.

Task characteristics may affect how a party manages conflict. When performing competitive tasks, an individual may be more likely to either compete or avoid. However, when performing a cooperative task within a group setting, collaborating and compromising are more likely to be used to manage within-group conflict. In addition, if the individual is performing a task that is highly interdependent with the other party's task, he or she is more likely to prefer a collaboration strategy.

In addition to these, organizational culture and norms are likely to affect both parties' behaviors during conflict. For instance, some conflict management strategies may be taboo within an organization. Rules and procedures are likely to affect conflict management as well because these can constrain certain behaviors while requiring other behaviors.

Structural conditions are considered fixed in the short run during the conflict process. However, in the long

run, structural conditions may be manipulated to make the conflict process more effective for the parties and the organization of which they are members. Some structural conditions can be affected by individual or organizational intervention. That is, parties can make adjustments to these conditions to influence how conflict is managed when it occurs. For instance, if increased stress results in undesirable behaviors and outcomes of a conflict episode, an organization may benefit by fostering an environment of low stress for its members. Intervening in this way can improve the way in which conflict is managed throughout the organization. Understanding individual and organizational characteristics that influence conflict can help managers and organizations to deal with conflict more effectively.

5. CONFLICT OUTCOMES

During a conflict episode, the primary parties and any third parties are continually working toward some sort of resolution. The ideal or acceptable outcomes of a conflict episode are likely to be related to the type of conflict management strategy employed by a primary party or to the type of intervention used by a third party. However, a conflict episode may have outcomes that go beyond what is expected by the parties involved. Conflict has several distinct types of outcomes: task, performance, and structural. Task outcomes include the resolution of the issue and the direct repercussions of the specific episode. Conflict can also have an effect on individual or group performance, for better or for worse, in the case where the conflict occurs in an organizational context; thus, performance is related to the direct repercussions portion of task outcomes. Finally, the effects of the episode on structural characteristics are a third distinct outcome. Thus, the three categories of outcomes that are discussed here are issue resolution, performance outcomes, and structural characteristics.

Issue resolution outcomes are related to the types of strategies employed to manage conflict. Much of the negotiation literature suggests that an agreement can result in either a distributive outcome or an integrative outcome. In addition, a third possibility is an impasse, that is, no agreement at all. A distributive outcome is one in which there is a fixed resolution that is either taken wholly by one party or somehow divided between the parties. An integrative outcome, on the other hand, is one in which the parties create an arrangement in which both parties are able to fully or significantly achieve their goals in resolving the conflict. This seems to be an ideal outcome, but in some situations integrative outcomes are not possible. In such cases, impasses and distributive outcomes are the only possibilities.

Jehn has contributed greatly to the research on performance outcomes of interpersonal conflict in organizations. Increased or decreased performance is a conflict outcome of which the primary parties might be unaware. That is, during conflict or as a result of conflict, each party's performance in an organization may either suffer or benefit. Two types of conflict with differing effects on performance in organizations are task conflict and relational conflict. Task conflict is concerned with the content and goals of the work itself such as a disagreement between managers on which market to enter next. Relational conflict focuses on the interpersonal relationship between the parties such as one party having a personal vendetta against the other party.

Moderate to high levels of task conflict appear to result in better performance in an organization. This is likely due to the fact that when bad or less effective ideas are opposed, countering with good or more effective ideas is likely to have a positive effect on performance. On the other hand, relational conflict has a negative effect on an individual's performance, probably due to the distractive effects of interpersonal feuding. In addition, each type of conflict is unique, with one type of conflict rarely transforming into the other (e.g., task conflict becoming relational conflict). This suggests two important points. First, an individual's work performance is affected by conflict. Second, whether an individual's work performance is affected positively or negatively by conflict depends on whether the conflict faced is task conflict or relational conflict.

When a conflict episode ends, the effects of the episode can affect future conflict episodes by influencing or changing characteristics of the parties or the organization. For instance, in 2000, Frone found that conflict with a supervisor can cause a worker to be less committed to an organization and can lead to lower job satisfaction and increase a worker's intention to quit. Similarly, a particularly influential conflict episode can lead to instituting or changing organizational policies on how conflict is handled such as creating an "open door" policy and requiring managers to oversee conflicts between coworkers. Thus, one unintended outcome of a conflict episode might be changes in individual or organizational characteristics that may linger and have important effects on future conflict episodes involving those parties or within that organization.

This suggests that conflict outcomes can be intended or unintended, direct or indirect, and beneficial or detrimental. It is likely that some of the outcomes of a conflict episode will be unintended such as performance outcomes and structural changes. The intended outcome, usually the issue resolution, may be the primary or only focus of a party engaged in a conflict episode. However, it is important for the two parties to recognize that conflict can lead to unintended detrimental outcomes that can outlast the intended beneficial outcomes they are trying to achieve. It is even more important for an organization to recognize that internal interpersonal conflict can result in detrimental outcomes for the firm, requiring third parties to intervene on behalf of the organization's interests and for the benefit of the two parties.

6. THIRD-PARTY INTERVENTION

In some situations, the two parties involved in conflict need assistance in reaching a resolution. For instance, if each party refuses to make any concession and neither party can exert influence over the other party, the result is an impasse. If a resolution is necessary, a third party may get involved to assist in resolution. In other instances, resolution of the conflict may have an effect on organizational interests and a third party may get involved to look out for those interests. Because third parties are often involved in managing conflict, it is important to understand the ways in which third parties can help (or hinder) the management of conflict. It is also helpful to understand what factors affect how third parties will intervene and the effects of such intervention on the organization and the primary parties.

6.1. Methods of Intervention

Two aspects of a conflict situation are relevant to third-party intervention: the conflict process and the resolution or outcome of the conflict episode. Process intervention is primarily concerned with managing the manner in which the primary parties interact. Outcome intervention is a matter of whether the third party has any control over the final decision regarding resolution. In 1985, Lewicki and Sheppard presented a set of third-party intervention methods, each of which exerts high or low process or outcome control. These intervention methods can be seen in Fig. 2.

Using high outcome and process control is referred to as inquisitorial intervention and has been shown

FIGURE 2   Third-party intervention methods. Reprinted from Lewicki, R. J., & Sheppard, B. H. (1985). Choosing how to intervene: Factors affecting the use of process and outcome control in third party dispute resolution. *Journal of Occupational Behavior, 6,* 49–64. Copyright John Wiley & Sons Limited. Reproduced with permission.

empirically to be the most common form of intervention employed by managers in dealing with conflict between subordinates. Using this type of intervention, a third party is likely to direct the interaction between the primary parties, deciding what information should be shared, deciding what is relevant or irrelevant, and possibly instructing parties as to how they should behave during the process. The third party is then likely to decide on a good resolution independently and to enforce this on the primary parties.

High outcome and low process control by a third party is referred to as adversarial intervention and is also a common method of intervention used by managers. This method is compared to the American judicial system in which parties present arguments and evidence with little or no direction from the third party. After each party has made a case, the third party makes the decision of how the conflict will be resolved and may even enforce the decision.

A less popular yet still used method of third-party intervention is the low outcome and low process control technique providing impetus. When using this method, a manager may listen to the primary parties to get a basic understanding of what the conflict is about and then suggest that the parties figure out a resolution themselves quickly or else some sort of punishment will be given or enforced. This is similar to a parent telling two children that they had better work out their differences or else they will be punished. A style with similar control as this method, yet more constructive and less threatening, is a method in

which the third party listens to disputants' views, incorporates their ideas, and asks for resolution proposals from the primary parties.

The low outcome and high process control method, or mediation, has been normatively described as the ideal means of third-party intervention but is rarely, if ever, used by managers in handling conflict between subordinates. It is suspected that this type of intervention is not used due to the large time investment required to direct the conflict process yet allow the primary parties to decide on a mutually beneficial resolution. Managers may prefer the other three methods of intervention so as to bring the conflict to a swift conclusion.

A third party is more likely to exert greater outcome control when time pressure is high, when the two parties are not expected to interact frequently in the future, and when the conflict outcome will have an impact on other parties and/or the organization. Greater process control is more likely to be used when the parties are less likely to interact in the future. Thus, it is suggested that managers are likely to use inquisitorial intervention when there are disputes between subordinates who do not interact very often. In addition, managers as third parties are more likely to use more autocratic (high-control) methods than are peers as third parties.

6.2. Outcomes of Intervention

How does one determine to what degree third-party intervention has succeeded? The effectiveness of a third party's involvement can be interpreted according to the achievement and quality of a resolution or according to the primary parties' perceived fairness of the process and outcome. Much of the research on the effects of third-party intervention focuses on the latter. Because third parties can vary in the amount of control exerted over the conflict process and outcome, the primary parties can have different perceptions about the fairness of the process and the outcome.

Procedural justice refers to the degree to which a party perceives that the conflict process was handled fairly by a third party. A party is more likely to perceive the process as fair if a compromise is reached or if the outcome is otherwise favorable to that party. In addition, a party is likely to perceive greater procedural justice when a third party uses mediation, which is a more facilitating role. Providing impetus and inquisitorial intervention may lead to negative perceptions of fairness in the conflict process because these roles are more autocratic.

Distributive justice refers to the degree to which a party perceives the outcome of an episode as fair. In a similar reflection of procedural justice, a party is more likely to perceive the outcome as fair if a compromise is reached or if the outcome is favorable to that party. Therefore, it is no surprise that a party is more likely to perceive an outcome as unfair when the outcome favors the other party. Unlike procedural justice, however, the method of intervention does not appear to have an impact on perceived fairness of the outcome. Research has shown that a party is more likely to perceive the outcome as fair if the process is perceived as fair.

Another possible outcome of a conflict episode is the lack of reaching a resolution, that is, an impasse. The use of autocratic intervention by third parties often increases the chances of impasses or one-sided resolutions, whereas facilitating methods of intervention often increase the chances of compromise resolutions. In addition, an impasse is more likely if a peer, rather than a manager, takes on an autocratic third-party role.

7. CONCLUSION

Interpersonal conflict is an unavoidable aspect of organizational life and can have a substantial impact on individuals and organizations. As noted in this article, conflict can affect task outcomes, performance, and organizational structure. Thus, learning to manage conflict is a worthwhile investment for both individuals and organizations.

This article summarized information that may help to develop a general understanding of interpersonal conflict in organizations. The conflict process was first described in an effort to elucidate the sequence of a conflict episode. Five conflict management strategies were discussed along with suggestions as to when each might be more appropriate. Individual and organizational characteristics that may influence both the development and management of a conflict episode were presented. The effects or outcomes of conflict were noted, as were the possibilities for third-party intervention. Because interpersonal conflict is inevitable, the challenge for individuals in organizations is to use this information to manage conflict more effectively such that positive outcomes are obtained.

See Also the Following Articles

Industrial/Organizational Psychology, Overview ■ Organizational Culture and Climate

Further Reading

Antonioni, D. (1998). Relationship between the Big Five personality factors and conflict management styles. *International Journal of Conflict Management*, 9, 336–355.

Blum, M. W., & Wall, J. A., Jr. (1997, May–June). HRM: Managing conflicts in the firm. *Business Horizons*, pp. 84–87.

Brockmann, E. (1996). Removing the paradox of conflict from group decisions. *Academy of Management Executive*, 10, 61–62.

Frone, M. R. (2000). Interpersonal conflict at work and psychological outcomes: Testing a model among young workers. *Journal of Occupational Health Psychology*, 5, 246–255.

Jehn, K. A. (1997). A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative Science Quarterly*, 42, 530–557.

Karambayya, R., & Brett, J. M. (1989). Managers handling disputes: Third-party roles and perceptions of fairness. *Academy of Management Journal*, 32, 687–704.

Karambayya, R., Brett, J. M., & Lytle, A. (1992). Effects of formal authority and experience on third-party roles, outcomes, and perceptions of fairness. *Academy of Management Journal*, 35, 426–438.

Lewicki, R. J., & Sheppard, B. H. (1985). Choosing how to intervene: Factors affecting the use of process and outcome control in third party dispute resolution. *Journal of Occupational Behavior*, 6, 49–64.

Lynch, D. (1997, May). Unresolved conflicts affect bottom line. *HRMagazine*, pp. 49–50.

Moberg, P. J. (2001). Linking conflict strategy to the Five-Factor Model: Theoretical and empirical foundations. *International Journal of Conflict Management*, 12, 47–68.

Mushref, M. A. (2002, November). Managing conflict in a changing environment. *Management Services: Journal of the Institute of Management Services*, pp. 8–11.

Oetzel, J. G. (1999). The influence of situational features on perceived conflict styles and self-construals in work groups. *International Journal of Intercultural Relations*, 23, 679–695.

Pondy, L. R. (1967). Organizational conflict: Concepts and models. *Administrative Science Quarterly*, 12, 296–320.

Pruitt, D. G. (2001). Conflict and conflict resolution, social psychology of. In N. J. Smelser and P. B. Baltes (Eds.) *International encyclopedia of the social and behavioral sciences*. San Diego: Elsevier Science.

Rahim, M. A., Buntzman, G. F., & White, D. (1999). An empirical study of the stages of moral development and conflict management styles. *International Journal of Conflict Management*, 10, 154–171.

Thomas, K. W. (1992). Conflict and negotiation processes in organizations. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3, pp. 651–717). Palo Alto, CA: Consulting Psychologists Press.

Van Slyke, E. J. (1999, November). Resolve conflict, boost creativity. *HRMagazine*, pp. 132–137.

Wall, J. A., Jr., & Callister, R. R. (1995). Conflict and its management. *Journal of Management*, 21, 515–558.

Conformity across Cultures

Rod Bond

University of Sussex, Brighton, United Kingdom

GLOSSARY

collectivist cultures Cultures in which the emphasis is on interdependence and the self is defined in relation to group membership.

cultural group A group sharing a common culture.

developed society A society with developed systems of political authority and use of technology in economic production.

individualist cultures Cultures in which individuals are independent of one another and social behavior is largely determined by personal goals rather than collective goals.

pastoral society A society in which grazing sheep and/or raising livestock form the primary mode of subsistence.

subsistence economy An economy that meets little more than the basic necessities and is typically characterized by low levels of production used for consumption rather than exchange.

Theories of conformity have long stressed the importance of cultural values in shaping people's responses to group pressure. Asch, for example, believed that his classic demonstrations of conformity—of people's willingness to "call White Black"—raised questions about the values fostered by American society and about its educational practices. Others have argued that the early demonstrations of conformity in the United States, especially those of Asch, reflect the particular climate of the McCarthy era and would not be found in the more tolerant culture of post-cold war society. In line with these arguments, evidence for the cultural roots of conformity has come from two types of comparisons: those across cultures and those within a culture at different periods in its history.

1. COMPARISONS ACROSS CULTURES

Studies that have compared conformity across cultural groups can be divided into three broad types: comparisons of subsistence economies, comparisons of developed societies, and comparisons of different cultural groups within a society.

1.1. Comparisons of Subsistence Economies

Berry has argued that different modes of subsistence require different human qualities that are reflected in cultural values and socialization practices and that give rise to differences in conformity. It has been found that pastoral and agricultural peoples emphasize obedience and responsibility in their socialization practices because these economies need people who are

conscientious and compliant, whereas hunting and fishing economies emphasize independence, self-reliance, and individual achievement because these economies need people who are individualist and assertive. Berry argued that if this is the case, pastoral and agricultural peoples should also be more likely to conform than are hunting and fishing peoples. Consistent with this reasoning, he found greater conformity on a line judgment task among the Temne of Sierra Leone, an agricultural society with strict disciplinarian socialization practices, than among the Eskimo of Baffin Island, a hunting and fishing society whose socialization practices are lenient and encourage individualism. However, his later research in Australia and New Guinea, as well as among North American Indians, obtained weaker support for this theory. He also found that where a subsistence economy had been exposed to Western values, there was less cross-cultural variation in conformity.

1.2. Comparisons of Developed Societies

Comparisons of developed societies have yielded a mixed set of findings. Several studies have reported cross-cultural differences in conformity that are explained by the relative values attached to conformity in the societies concerned. One of the earliest comparisons by Milgram found greater conformity on a perceptual task among Norwegian students than among French students, and this was taken as reflecting the fact that Norway is less tolerant of deviance than is France. Some investigators have replicated the Asch experiment, in which participants judge which of three lines is the same length as another line after hearing the unanimous, but manifestly erroneous, judgment of a group of other participants who are in fact confederates of the experimenter. Replications in Brazil, Zimbabwe, Ghana, and Fiji, as well as among the Chinese, found high levels of conformity when compared with the original results, and these findings were interpreted as reflecting the higher value attached to conformity in these societies.

Anticipated differences have not always emerged, however, and further studies of conformity from countries as diverse as Brazil, Lebanon, Zaire, France, Kuwait, Portugal, and Hong Kong all have reported comparable results, leading some authors to remark on the cross-cultural stability of the classic demonstrations of conformity. Studies in Japan, a country that at

first sight might be expected to have higher levels of conformity, have also been inconclusive and point to the importance of taking into account the relationship between the individual and the group. When the majority are strangers to the individual, there is some evidence for a lower level of conformity (relative to the United States) among Japanese students, but other evidence suggests a higher level of conformity when the majority are friends of the individual.

1.3. Comparisons of Cultural Groups within a Society

Some investigators have compared cultural groups within a society as a way in which to study the cultural roots of conformity, although this has not proved to be a fruitful line of inquiry. There is some indication that conformity is higher among minority groups than among majority groups. For example, some studies have found higher conformity among Blacks than among Whites in the United States, and another study found higher conformity among Puerto Ricans than among Whites. A study in Britain found higher conformity among West Indians than among Whites. A study in Fiji found greater conformity among Fijian Indians than among native Fijians. However, there is no consistent pattern of results, nor is it clear why such differences have been found.

1.4. Conformity and Individualism–Collectivism

Cross-cultural comparisons have yielded a mixed set of findings. This may be due, in part, to the methodological limitations of many of the studies in that frequently poor control is exercised over differences in sample and method between the cultures that are compared. Also, comparisons typically are made with single samples from each culture, and evidence from other studies carried out in a culture is disregarded. The explanation for cross-cultural differences, when they are found, is usually post hoc, and there has been little attempt to directly assess the values presumed to mediate responses to group influence. To address these concerns, Bond and Smith in 1996 conducted a meta-analysis of 133 Asch-type conformity studies from 17 countries that was able to provide tighter control over extraneous factors and combine all available evidence from various cultures. Moreover, they proposed that cross-cultural differences in conformity would depend on the extent to which a culture subscribed to

collectivist values as opposed to individualist values. Individualism–collectivism has emerged as the most important dimension on which cross-cultural differences in values have been described, albeit one that continues to be a central topic of debate. In individualist cultures such as European Americans, individuals are independent of one another and social behavior is largely determined by personal goals rather than collective goals. Where personal goals conflict with collective goals, it is personal goals that tend to be prioritized. In contrast, in collectivist cultures such as East Asian societies, the emphasis is on interdependence and the self is defined in relation to group membership. Social behavior is largely determined by collective rather than personal goals, and collective goals are given priority whenever they are in conflict with personal goals. Conformity to the majority would be more likely in collectivist cultures than in individualist cultures given the higher value placed on maintaining harmony in group relations. This was what Bond and Smith found. Using measures of collectivist values derived from three independent cross-national value surveys, they found that studies conducted in collectivist countries tended to report larger conformity effects than did those conducted in individualist countries.

2. COMPARISONS WITHIN A CULTURE AT DIFFERENT PERIODS IN ITS HISTORY

A second line of evidence for the cultural roots of conformity comes from the observation that replications of conformity experiments within a society produce different results at different points in time. Larsen conducted three replications of the Asch experiment and interpreted the different levels of conformity as reflecting sociopolitical changes in American society. The more questioning attitude of students of the Vietnam War era explained the lower level of conformity found in 1974, whereas the decline of student activism explained the rise of conformity levels by 1979. In 1988, conformity had declined again somewhat, and this was seen as reflecting the increase in protest activities. Similarly, Perrin and Spencer were unable to replicate the Asch experiment with British students in 1980, and they attributed this to their more questioning attitude compared with Asch's respondents during the 1950s. They concluded that the Asch experiment was not an enduring social

psychological phenomenon but rather one rooted in its historical context. A few years later, another study found some conformity among British students that was attributed to the greater cohesiveness brought about by the Falklands war.

Not all investigators have found it difficult to replicate the Asch experiment, however, and several other recent studies of university students conducted in Britain, Belgium, Holland, and Portugal all have found reasonably high levels of conformity comparable to those found in Asch's research. Another line of argument, proposed by Lamb and Alsifaki, is that levels of conformity have been steadily on the increase, drawing on the hypothesis that modern industrial societies are characterized by increasing numbers of "other-directed" types who are more easily influenced by peer pressure.

This line of inquiry is highly speculative given that there are many potential reasons, apart from changes in the sociopolitical context, why two studies conducted at different points in time would give different results. Nevertheless, Bond and Smith's meta-analysis of Asch-type conformity studies enabled a synthesis of a number of studies and controlled for differences in study design and method. Using only those studies carried out in the United States, of which there were 97 in total, Bond and Smith found that the level of conformity was related to the time when the study was carried out. Studies carried out more recently have tended to report lower levels of conformity than did earlier studies, perhaps suggesting that there has been a greater value placed on independence in U.S. society during the post-World War II era.

3. CONFORMITY OR HARMONY?

The social psychological literature, by and large, portrays conformity in negative terms. The individual is pitted against the group, and to yield to group pressure is an act of submission and a sign of weakness. Individuals are expected to express their own opinions, and those who continue to do so in spite of pressures to do otherwise are admired. Such a view of conformity is rooted in Western individualist values in which self-expression and independence from the group are highly valued. Viewed from the standpoint of collectivist values, however, such behavior may seem selfish and immature. In such societies, individuals should seek to maintain harmony with the group, and open disagreement can be a source of embarrassment. Thus, the same behavior might be viewed as conformity in an

individualist culture (and be judged negatively) but be viewed as maintaining harmony in a collectivist culture (and be judged positively).

In 1999, Kim and Markus demonstrated the different values placed on conformity and uniqueness in collectivist and individualist cultures. For example, they found that magazine advertisements in Korea were more likely to have conformity themes, emphasizing respect for collective values and beliefs as well as harmony with group norms, than were those in the United States. In contrast, advertisements in the United States were more likely to have uniqueness themes consisting of freedom, choice, individual uniqueness, and rebelling against collective values and beliefs. These authors also studied preferences for unique or common elements of abstract geometrical figures and found that European Americans preferred unique elements, whereas East Asians preferred common elements. Thus, conformity takes on different meanings depending on the network of values and beliefs that comprise the cultural context within which the behavior occurs, and these differences in meaning in turn will entail that conformity or independence will carry different affective consequences.

4. CONCLUSION

The literature on cross-cultural differences in conformity has provided a mixed set of results. This is due, in part, to the difficulty in ensuring comparable methods and samples when conducting cross-cultural research, and conformity studies are in any case difficult to carry out because it is so easy to arouse suspicion among participants. Most studies have used the classic conformity paradigms of Asch or Crutchfield and have used mostly perceptual judgment tasks. The major limitation has been the lack of attention to underlying process or theoretical elaboration of relevant features of the situation. Recent work relating individualism–collectivism to conformity goes some way in redressing

this issue and suggests that people in collectivist cultures will conform more. There is also some evidence that conformity in U.S. society has declined since World War II. However, no studies have yet directly assessed how collectivist values relate to conformity. Moreover, the nature of collectivism suggests that the relation of the individual to the group is crucial and that it is not generally true that people in collectivist societies will conform more. In collectivist cultures, individuals will seek to maintain harmony with in-group members but may be less concerned about maintaining harmony with out-group members. Finally, care must be taken in interpreting behavior within its cultural context. Conformity may reflect a different capacity to resist group pressure or a different determination to avoid embarrassment. Similar behavior may take on different meanings and reflect different processes in different cultural contexts.

See Also the Following Articles

Acculturation ■ Cross-Cultural Psychology, Overview ■ Cultural Complexity

Further Reading

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, *119*, 111–137.

Furnham, A. (1984). Studies of cross-cultural conformity: A brief and critical review. *Psychologia*, *27*, 65–72.

Kim, H., & Markus, H. (1999). Deviance or uniqueness, harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology*, *77*, 785–800.

Mann, L. (1988). Culture and conformity. In M. H. Bond (Ed.), *The cross-cultural challenge to social psychology* (pp. 184–187). Newbury Park, CA: Sage.

Smith, P. B., & Bond, M. H. (1998). *Social psychology across cultures: Analysis and perspective* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Connectionism

Yoshihisa Kashima

University of Melbourne, Melbourne, Australia

GLOSSARY

delta rule (backpropagation rule) A well-known algorithm used in a connectionist network for adjusting the strength of links among artificial neurons as a function of the amount of errors generated when the network predicts outputs based on its inputs.

distributed approach to meaning An approach where a meaningful concept is represented in a connectionist network by a pattern of activation over a number of nodes or artificial neurons.

"exclusive or" operation A logical operation where either one or the other alternative is selected and where both alternatives may not be selected.

feedforward network A type of connectionist network architecture in which input signals are transmitted forward, but not backward, to output nodes.

generalized delta rule An algorithm for adjusting the strength of links among artificial neurons so that a connectionist network can learn input–output relations.

Hebbian learning rule An algorithm for adjusting the strength of connections among artificial neurons where the amount of change in connection strength is a function of the amount of activation of the connected neurons; it is named after Donald Hebb, who first proposed the idea for the learning involving real neurons.

localist approach to meaning An approach where a meaningful concept is represented in a connectionist network by a corresponding node or artificial neuron.

multilayer feedforward network A type of connectionist network architecture in which several clusters (or layers) of artificial neurons are connected with each other and input signals are transmitted only forward through multiple layers of artificial neurons to an output layer.

recurrent network A type of connectionist network architecture in which input signals are transmitted forward from input nodes to output nodes but also backward from output nodes to input nodes.

sigmoid function A mathematical function that links input amount to output amount (typically of an artificial neuron) where output increases slowly when the amount of input is small or large but increases rapidly when the input amount is medium.

supervised learning A type of learning that occurs in connectionist networks where outputs that are to be associated with inputs are provided as standards or "teachers" from which to learn.

synaptic link A link between neurons in a nervous system through which neural signals are transmitted.

tensor network A type of connectionist network architecture in which several clusters of artificial neurons are connected with each other.

threshold function A mathematical function that links input amount to output amount (typically of an artificial neuron)

where an output occurs when an input reaches a certain level or a threshold.

unsupervised learning A type of learning that occurs in connectionist networks where outputs that are to be associated with inputs do not act as standards or "teachers."

Connectionism is a way of conceptualizing psychological processes that gets its inspiration from the way in which the biological brain processes information. In this approach, psychological phenomena are explained in terms of activations of interconnected neuron-like units whose operations are governed by simple rules.

1. INTRODUCTION

Although ideas akin to connectionism may be attributed to past thinkers such as Aristotle and William James, modern connectionism began its tenure during the mid-20th century. Alongside the theoretical and actuarial development of digital computing based on von Neumann's concept of serial computing (most current computers are of this type), there emerged a research tradition that took seriously the biological brain as a metaphor for information processing. In this view, just as the brain does its work through neurons and synaptic links, the mind too may be conceptualized as functioning through neuron-like simple units and connections among them. Connectionism, also known as neural networks, is so called due to the network of connections among the neuron-like units that are postulated to explain psychological phenomena.

McCulloch and Pitts, Hebb, and Rosenblatt conducted their pioneering work during the 1940s and 1950s. Nevertheless, Minsky and Papert pointed out in 1969 that some of the simple logical operations such as "exclusive or" cannot be performed by the then available mechanisms. These limitations, together with the development of serial computers as a main technological tool of computation, resulted in a decline in the research on brain-inspired information processing mechanisms. While Anderson, Fukushima, Grossberg, Kohonen, and others continued to work during the 1960s and 1970s, it was the 1980s that witnessed the renaissance of connectionism. A mechanism called the generalized delta rule, which rectified connectionism's problem identified by Minsky and Papert, was developed. A network of simple neuron-like units was shown to be able to model psychological processes such as perception, memory, and learning. Rumelhart and colleagues'

Parallel Distributed Processing, published in 1986, popularized the connectionist approach.

By this time, however, the approach that used the von Neumann serial computer as a metaphor of mind (a psychological process is considered to be a series of discrete steps of mental operations on some mental representations) had established itself as the dominant research program in psychology and cognitive science. Controversies ensued as to how connectionism differs from the serial processing approach, which approach explains psychological processes better, and whether connectionism provides a realistic model of brain processes.

2. GENERAL FEATURES OF THE CONNECTIONIST COGNITIVE ARCHITECTURE

Although connectionist architectures vary considerably, they can be characterized in general terms. First, a connectionist network typically consists of a large number of units (sometimes called nodes) that are connected to each other with varying strengths (Fig. 1). The operation of each unit is governed by a simple rule (discussed later) that determines its activation state ("firing" or "resting") as a function of the inputs it receives from the connected units. The connection strength (often called weight) of a given unit to another unit indicates the amount of influence the first unit has over the second unit. Second, the pattern of activation of the units changes over time. The rules used in connectionist networks specify how the activation of a unit at one point in time (t) affects other units at a later time ($t + \Delta t$). Interunit effects occur in parallel, so that during a given time interval (Δt), the effects of all units on other units (a unit may have an effect on itself) take place. As
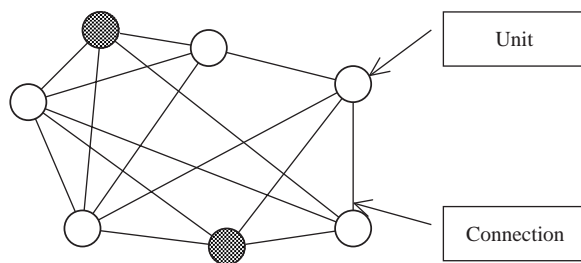

FIGURE 1   A hypothetical connectionist network. The shaded unit is activated.

a result, the state of a connectionist network can continue to change, potentially forever. However, under some circumstances, the state of a connectionist network shows a discernible temporal pattern. For instance, it may come to a stable pattern, oscillate between two or more patterns, or exhibit a chaotic pattern. Third, a connectionist network performs its function collectively. If one unit is activated, its activation spreads via its connections to other units in accordance with a rule. As these activations reverberate through the network, the network as a whole may then perform various functions such as the encoding and retrieval of information.

Connectionism differs from models of cognition based on the serial computer metaphor in several respects. A serial computer typically has a central processing unit (CPU), data, and procedures (i.e., software) that operate on the data, and the CPU executes the procedures to handle the data. One fundamental assumption of the serial computer metaphor of cognitive processes is that both data and procedures are written in terms of physical symbols. Physical symbols (e.g., the word "cat") refer to or stand for some objects and events in the world (e.g., the cat on the mat) and are themselves physically coded on a physical medium such as an ink mark on paper, an electronic circuit, or a synaptic circuit in the brain. In contrast, connectionism does not assume the existence of a single CPU; rather, it assumes the existence of a multitude of simple processing units. What the CPU does in a serial computer is governed by a simple or complex set of procedures that can be programmed (e.g., a simple calculator, complex word processing software), but processing units in connectionist networks have a simple and fixed rule. In other words, connectionist networks are not programmable in the same way in which a serial computer can be programmed—although they can be trained or taught (discussed later). Finally, connectionism does not assume the explicit coding of information in physical symbols. Despite these differences, both metaphors have been fruitfully used to advance psychological science. One metaphor appears to be more suited to model some psychological processes than to model others, but both metaphors often do a similarly good job of explaining a set of phenomena. At this point in history, they may be best considered to be different research programs with different theoretical assumptions and conventions rather than two competing paradigms.

Although neurally inspired, connectionist networks are not models of brain functions. Put differently, the relationship between the connectionist network and the brain is metaphorical rather than literal. To be sure, connectionist modelers often make use of design principles that are observable in the brain. For instance, just as the firing of a neuron is not a simple linear function of the excitation it receives, the activation of a unit in a neural network is often governed by a nonlinear function of the inputs it receives. Just as synaptic connections may be excitatory or inhibitory, connections among connectionist units may be excitatory or inhibitory. However, a learning rule such as the generalized delta rule, which plays an important role in neural network research, has no known corresponding mechanism in the brain. More generally, the brain anatomy is highly structured, but most neural networks do not have differentiated architectures that resemble the brain structure. In the long run, neural network research may become more informed by neuroscience (and, in fact, there is research that tries to model brain functions). Currently, however, connectionist networks are at best highly abstracted and stylized models of how the brain works. For this reason, connectionist networks are often called artificial neural networks.

3. LOCALIST AND DISTRIBUTED REPRESENTATIONAL SYSTEMS AND THEIR PROPERTIES

How do we understand objects and events in the world, and how do we experience our world as meaningful? Meaning is one of the most important questions for models of cognition. In cognitive models based on the serial computer metaphor, symbols gain their referential meanings by designation; a symbol is designated to refer to a certain object or event in the world. In this sense, meaning is built into the system of cognition. In neural networks, meaning is handled in two different ways. The localist approach is akin to the serial models of cognition. Here, a unit in a connectionist network is designated to have its referent. When this unit is activated, the person is considered to be in some sense "thinking about" the referent that the unit represents. In contrast, in the distributed approach, any given unit might not have a clear referent, and its activation cannot be interpreted as corresponding to any meaningful thought in and of itself. Rather, it is a pattern of activations of many units that imply a meaning. In other words, meaning is not localized to a unit but rather distributed across units. A unit in a distributed system then represents some microfeature of a meaningful

concept that may or may not be interpretable using symbols such as words and phrases. In this sense, a distributed approach models subsymbolic processes.

An advantage of the localist approach is its ease of interpretation. The activation of a unit means that the concept represented by the unit is being thought about; changes in activation of a network can be easily interpreted as changes in the content of thought. Although the localist approach may be sufficient to model many psychological processes, the distributed approach might be necessary when some of the following features are important. First, the distributed representational system has a built-in capacity to model similarity. Because a concept is represented by a pattern of unit activation in the distributed system, the similarity between one pattern and another pattern of activation can be naturally interpretable as a similarity in meaning. Second, the capacity to represent similarity enables the distributed system to have the capacity for psychological generalization as an inherent aspect of the system. If a system responds to a pattern in a certain way, it responds to a similar pattern in a similar way; that is, the system generalizes as a natural consequence of its distributed format. Third, these capacities often result in a graceful degradation if the system is damaged. For instance, if one unit in a localist network is damaged or removed, it loses the capacity to represent the corresponding concept. However, even if one unit in a distributed network is lost, it still retains most of the capacity to represent concepts even though the clarity of representation might be compromised to some extent.

FIGURE 2 Information processing unit and activation functions.

4. DYNAMICS OF AN INFORMATION PROCESSING UNIT

In a time interval (Δt), an information processing unit in a neural network receives inputs from other units, computes its net inputs, gets activated according to an activation function, and sends out its output to other units according to an output function. Figure 2 shows a schematic picture in which Unit i receives inputs from its neighbors, Unit j and other connected units, and sends out its output. When a neighboring Unit j sends out its output, o_j, it spreads via the connection with its connection strength or weight, w_{ij}, to Unit i. The input from j to i then is the product of the output and weight, that is, $w_{ij}o_j$. The inputs from other units are also computed in a similar fashion. The Unit i then sums all inputs to compute its net input, that is, $net_i = \Sigma w_{ij}o_j$, where the summation is over j, which indexes units that send out their outputs to Unit i.

Whether, or how much, Unit i is activated depends on the net input and the activation function of the unit. Typically, the resting state of a unit is indicated by the activation level of zero and the firing state is 1; however, depending on the architecture, a unit may take a value between –1 and +1 or between negative and positive infinity. The activation function determines activation level of the unit as a function of net inputs. Although activation function may be a simple linear function, it often takes a nonlinear form. One example is a threshold function (lower left panel of Fig. 2), where the activation level remains zero if the net inputs remain lower than a threshold value but its level becomes 1 when the net inputs exceeds the threshold.

Another example is a sigmoid function, where activation changes smoothly as a function of net inputs in the shape depicted (lower right panel of Fig. 2). These nonlinear activation functions are critical in some applications. The output of Unit i is determined by the activation level of the unit and its output function. Although output is often identical to activation (i.e., $o_i = a_i$), it may vary for some applications.

When activation spreads in only one direction (i.e., from Unit j to Unit i), activation feeds forward from one set of units to another set in one time interval, whose activation then spreads to a third set in the next interval, and so on. This type of network, called a feedforward network, basically transforms one set of activations in the first set of units into another set of activations in the last set of units. Its main function is to convert one set of signals to another set of signals. In psychological terms, it may be useful, for instance, for modeling how a given situation or stimulus may induce a set of responses. However, activation may spread bidirectionally (from Unit j to Unit i as well as from Unit i to Unit j), and the connection strength in one direction may or may not be the same as that in the opposite direction. In these cases, more complex dynamics emerge where activation may spread back and forth and reverberate through the network. This type of network, sometimes called a recurrent network, may tend toward a stable pattern of activation in which activation levels of the units change little or oscillate among several stable patterns. These networks may be used to model memory retrieval processes. Recurrent networks may sometimes exhibit an even more complex pattern known as chaos.

5. DYNAMICS OF INTERUNIT CONNECTIONS

Interunit connections in a connectionist network may be fixed or modifiable. In the latter case, connections are modified in accordance with a learning rule, so that the network may learn contingencies among events in its environment. For illustrative purposes, let us consider a simple example of a distributed feedforward network with two layers. In this network, each layer has three units and there is no connection among the units within each layer, but all of the units in one layer are fully connected to the units in the other layer (Fig. 3). Let us suppose that the network is to learn that when Event A happens, Event B follows. Further
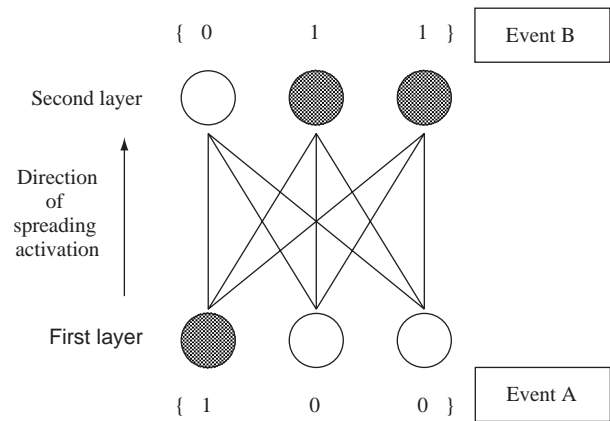

FIGURE 3 A two-layer distributed feedforward network. Shaded units are activated at +1. The bottom layer represents Event A, and the top layer represents Event B.

assume that Event A is encoded by the units in the first layer; when the first unit is activated (+1) but the other units are not activated (0), this means that Event A is represented in the network. Event B is encoded by the units in the second layer; when the first unit in this layer is not activated (0) but the second and third units are activated (+1), this means that Event B is represented in the network.

For the network to learn the association between Event A and Event B, the network needs to modify its connections so that when Event A is represented in the first layer, its activation would spread to the second layer and Event B should then be represented in the second layer. There are two basic types of learning rules that can accomplish this. One is called unsupervised learning, and the typical learning rule is called the Hebbian learning rule, named after the psychologist Donald Hebb. The rule is to strengthen the connection between the units that are activated at the same time. In this example, the connections between the first unit in the first layer and the second and third units in the second layer should be strengthened (e.g., from 0 to 1), whereas the other connections should remain unchanged (e.g., 0). In this case, assuming that input, activation, and output all are the same for a given unit, it is easy to see that the pattern of activation {1 0 0} in the first layer would spread to the second layer and activate the second and third units but would not activate the first unit in the second layer. More formally, this rule is more generally expressed by the following equation: $\Delta w_{ij} = \varepsilon a_i o_j$, where Δw_{ij} is the amount of change in connection strength from Unit j

to Unit *i*, a_i is the activation level of Unit *j*, o_j is the output of Unit *j*, and ε specifies a learning rate, that is, the parameter that governs the speed of learning.

The other type of learning is called supervised learning, for which the typical rule is the delta (or back-propagation) rule. Here, the connections are changed so that if Event A is represented in the first layer, it predicts Event B in the second layer. Event B acts as a "teacher" that is to be emulated by the network. Here is how the rule works. Assuming that all of the connection strengths are initially zero, all of the inputs to the units in the second layer would be zero. For the first unit in the second layer, then, the teacher is zero and the activation is also zero, that is, no error. When there is no error, the connection strengths remain the same. So, the connections coming into the first unit in the second layer remain zero. However, the teacher for the second and third units says that they should be activated (i.e., +1) even though they are not activated (i.e., 0). The difference between the teacher and the predicted activation is an error, which is then used to modify the connection strength. The rule is to change the connection strength to reduce the error so that the connection with the activated unit gets stronger but that with the resting unit remains the same. Therefore, the connection from the first unit in the first layer to the second and third units in the second layer is strengthened (e.g., from 0 to +1), but the other connections remain unchanged (e.g., 0). Again, this rule can modify the connection strengths appropriately so that the network predicts Event B when Event A is encoded. More generally, this rule is expressed by the following equation: $\Delta w_{ij} = \varepsilon(t_i - a_i)o_j$, where Δw_{ij} is the amount of change in connection strength from Unit *j* to Unit *i*, t_i is the teacher's activation level, a_i is the predicted activation level of Unit *j*, o_j is the output of Unit *j*, and ε is the learning rate.

At one level, supervised learning differs from unsupervised learning in the presence of a teacher or supervisor in the learning process. Supervised learning is sometimes said to be less realistic because people learn without being told what they have expected is wrong by how much. However, even with supervised learning rules, there is no need to postulate the explicit presence of a human teacher. The only necessary assumption is that one of the events to be associated is regarded as a teacher. In other words, the "teacher" may be a human, a machine, or nature. More significantly, supervised and unsupervised learning rules differ in terms of their conceptualization of learning. On the one hand, unsupervised learning conceptualizes learning as the observation and learning of the co-occurrence of events in the world.

On the other hand, supervised learning presupposes a learning agent that implicitly predicts what follows what (e.g., Event B follows Event A) and learns from mistakes. In this sense, supervised learning may be able to model adaptation. However, networks with supervised learning may actually be too adaptive in that they might forget previously learned associations—sometimes even catastrophically—when new associations are learned. Unsupervised learning does not have this problem most of the time. Both the Hebbian and delta rules can be generalized to more complex forms.

6. EXAMPLES OF CONNECTIONIST ARCHITECTURES

A tensor network is an example of the connectionist architecture that embodies a generalized form of the Hebbian learning rule (Fig. 4). In the example in the previous section, the network learned the association between two events, but a tensor network is designed to learn to associate three or more events (Fig. 4 depicts a network that associates three events). One cluster of units represents one type of event, the units in one cluster are connected to the units in other clusters, and the connection among units (the connection among three units in Fig. 4) changes as a function of the activation levels of those units. This type of network is called a tensor network because the learning rule can be mathematically described by a mathematical representation called tensor. When one of the clusters of units is activated, its activation spreads to other units, which are then activated to retrieve stored information. A tensor network is suitable for integrating multiple pieces of information into a structured representation. It can bind together different types of information such as color, time of blooming, and
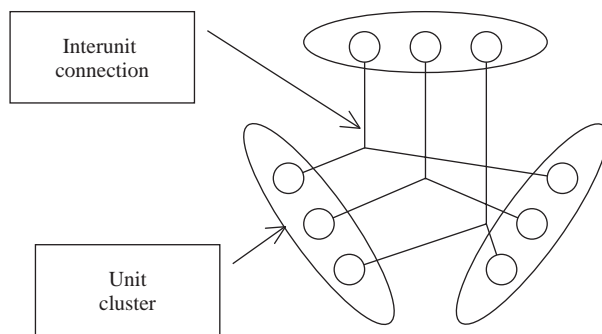

FIGURE 4 A tensor network that associates three events.

habitat; for instance, it can represent the name of a red flower that blooms during the early spring in wetlands. More generally, a tensor network can approximate the intersection of different sets.

A multilayer feedforward network is an example of the architecture that uses a generalized form of the delta rule (Fig. 5). In the previous example, there were two layers of units, but a multilayer network has three or more layers of units. The first and last layers are called input and output layers because they interface with the environment, encoding the incoming information (input layer) and decoding the outgoing response (output layer). The middle layers are called hidden layers because they are hidden from the environment. When the input layer units are activated, their activation spreads to the hidden layer units, which are then activated. The activation feeds forward so that the output layer units are activated. The output layer activation is compared to the relevant teacher, and the discrepancy (error) is then used to modify the connection strengths among the units. First, the connections between the output layer units and the units in the next last layer are modified, then the connections between the next last layer and the units before this layer (input layer in Fig. 5) are modified, and so on. Thus, the error information propagates backward, so to speak, to modify the connections. This is why the generalized delta rule is also called the backpropagation rule. Although a two-layer feedforward network cannot perform a simple logical operation of "exclusive or" (i.e., either/or), a three-layer network with a hidden layer can do so. More generally,

a hidden layer is useful for transforming an input representation to re-represent it.

Some architectures use the Hebbian or delta rule with further modifications to the basic network structure. In recurrent networks (sometimes called auto-associative networks), units may have connections that spread activation not only in one direction but also in the opposite direction. This type of architecture tends to amplify strong activation and dampen weak activation, sharpening the pattern of activation. It can be used for pattern completion; that is, when only a part of the original information is presented, this network can often fill in the missing part to complete the original pattern. Its practical uses include noise reduction and error correction.

Although these types of connectionist networks are based on somewhat different conceptions of learning and have different advantages and disadvantages, they may be combined to model more complex processes. A complex network that combines different types of network architectures is sometimes called a modular architecture. Each module is designed to perform a specified function, whose output is then used by another module for further processing. For example, a recurrent network may be used to clean noisy inputs so as to provide a clearer input to a tensor network, whose output may then be transferred to a multilayer feedforward network for modeling adaptive learning. Several modules may then function in tandem to provide a more complete model of human and animal psychological processes.
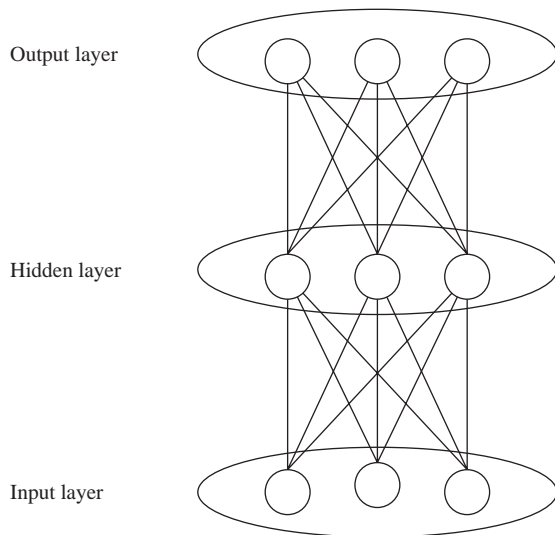
FIGURE 5   A three-layer feedforward network.

7. APPLICATIONS OF CONNECTIONISM

Connectionist networks provide versatile research tools for developing a model of human psychological processes. For example, they have been used to model some basic perceptual and cognitive processes such as pattern recognition, learning, memory, categorization, judgment and decision making, and natural language processing as well as more social psychological processes such as person impression formation, stereotyping, causal attribution, and stereotype formation and change. Some models have been used to discuss issues pertaining to developmental psychology. Some networks have been used to model neuropsychological issues; when parts of the interunit connections of a network are removed, the network exhibits behaviors that are similar to those of people with brain lesions. Other variants can model

animal behaviors as well. Although more applied areas of psychological research have been relatively slow to embrace connectionism, there are some signs of its use. Finally, the technology used in connectionist networks has been adopted not only to model psychological processes but also for engineering or other purposes. For instance, optic pattern recognition is useful for machine pattern recognition, and feedforward networks may be used for forecasting.

See Also the Following Articles

Cognitive Skills: Training, Maintenance, and Daily Usage ■ Cognitivism

Further Reading

Anderson, J. A., & Rosenfeld, E. (1988). *Neurocomputing.* Cambridge, MA: MIT Press.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1998). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Minsky, M., & Papert, S. (1969). *Perceptrons.* Cambridge, MA: MIT Press.

Read, S. J., & Miller, L. C. (1998). *Connectionist models of social reasoning and social behavior.* Mahwah, NJ: Lawrence Erlbaum.

Rumelhart, D. E., & McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing.* Cambridge, MA: MIT Press.

Conscientiousness

Robert R. McCrae

National Institutes of Health, Baltimore, Maryland, USA

GLOSSARY

cross-observer agreement Agreement between two or more observers on characteristics of a target person.

exvia In Cattell's system, a broad factor resembling introversion–extraversion.

heritable Determined in part by genes inherited from one's parents.

intercorrelate To vary together such that people who are high on one of the variables are likely to be high on the others as well.

orthogonal In factor analysis, referring to factors that are unrelated to each other.

psychopathology Mental and emotional disorders.

Conscientiousness is one of five major dimensions of individual differences in personality. It contrasts people who are methodical, purposeful, and deliberate with those who are disorganized, lazy, and hasty. Conscientiousness shows strong genetic influences. It is relatively low among adolescents and increases up to 30 years of age. Thereafter, mean levels and individual differences are generally stable throughout adulthood. Low levels of the Conscientiousness factor are associated with several forms of psychopathology, whereas high levels are associated with academic achievement and superior job performance.

1. CONSCIENTIOUSNESS AS A DIMENSION OF PERSONALITY

Personality traits are tendencies to think, feel, and act in consistent ways. People in every culture learn to describe themselves and others in trait terms, with words such as *nervous*, *energetic*, *original*, *altruistic*, and *careful* referring to personality traits. The English language has several thousand such words. In addition, personality psychologists working in a variety of theoretical perspectives have identified traits such as ego strength, exvia, tolerance of ambiguity, dependency, and socialization.

All of these traits are similar in some respects. All are normally distributed; that is, a few people score low, most score in the average range, and a few score high. All traits are relatively stable across time. All of them affect behavior, but their impact is often weak and observable only when averaged across a wide sample of occasions.

It has become clear during the past 20 years or so that nearly all personality traits are related to one or more of five basic personality dimensions or factors, most commonly labeled Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Together, these factors and their constituent traits are known as the Five-Factor Model (FFM). Each of the five factors is a broad disposition defined by many narrower

traits. Furthermore, the factors are orthogonal in that people who are very high in Conscientiousness may be introverted or extraverted and may be agreeable or antagonistic.

Many personality theorists have described traits akin to Conscientiousness. These include Tupes and Christal's dependability, Tellegen's constraint, Gough's norm adhering, Cloninger's persistence, Digman's will to achieve, and Cattell's control. These various names reflect somewhat different conceptions of the Conscientiousness factor, with some suggesting a proactive purposefulness and others suggesting an inhibitive self-control. In an attempt to represent the full range of relevant traits, Costa and McCrae in 1991 assessed six facets of Conscientiousness in their Revised NEO Personality Inventory (NEO-PI-R): competence, order, dutifulness, achievement striving, self-discipline, and deliberation. Factor analyses have shown that these six facets intercorrelate and define a clear factor in both men and women, as well as in both adults and children, in many different languages and cultures.

2. ORIGIN AND DEVELOPMENT OF CONSCIENTIOUSNESS

Terms such as *hard-working*, *reliable*, and *persevering* describe desirable aspects of character. Indeed, Conscientiousness was omitted from some personality systems because it was once believed to be a moral evaluation rather than a real psychological attribute. The reality of individual differences in Conscientiousness has now been clearly established by studies of cross-observer agreement. Peer and expert ratings confirm the self-reports that people make about their degrees of Conscientiousness. Furthermore, both self-reports and observer ratings of Conscientiousness predict real-life outcomes such as academic success.

During most of the 20th century, psychologists believed that personality traits could be divided into two categories: temperament and character. Temperament traits were thought to be biologically based, whereas character traits were thought to be learned either during childhood or throughout life. With the advent of the FFM, behavior geneticists began systematic studies of the full range of personality traits, and it soon became clear that all five factors are substantially heritable. Identical twins showed very similar personality traits even when they had been separated at birth and raised apart, and this was equally

true for both character traits and temperament traits. Parents and communities influence the ways in which Conscientiousness is expressed, but they apparently do not influence its level.

Currently, little is known about Conscientiousness in young children because the self-report inventories typically used to assess it are not appropriate for that age group. It is obvious, however, that there are individual differences on this factor at an early age. We know, for example, that some children have attention deficit/hyperactivity disorder, which is characterized in part by problems with concentration, organization, and persistence—traits related to Conscientiousness. Longitudinal and cross-sectional studies suggest that Conscientiousness is relatively low among adolescents but increases between 18 and 30 years of age. Through most of the rest of adult life, there is relatively little change in the average level of Conscientiousness, although it is less clear what happens during extreme old age. Individual differences are strongly preserved, meaning that a careful, neat, and scrupulous 30-year-old is likely to become a careful, neat, and scrupulous 80-year-old.

3. APPLICATIONS

It matters greatly whether one is purposeful, diligent, and persistent or aimless, undisciplined, and impulsive. Consequently, Conscientiousness is associated with a variety of important life outcomes.

3.1. Mental Health and Psychotherapy

Individuals who are low in Conscientiousness are at risk for a number of psychiatric disorders. Because they act without due consideration of the long-term consequences of their actions, and because they lack self-control, they are prone to alcohol and substance abuse. Because they are chronically low in motivation to pursue goals, they are more susceptible to demoralization and depression. Because they do not have a strong sense of social obligations, they are more likely to be psychopaths or to have antisocial personality disorder. In 1998, Ranseen and colleagues showed strong associations between Conscientiousness and adult attention deficit disorder.

In principle, very high levels of Conscientiousness might also lead to problems in living. Some theorists link high Conscientiousness to obsessive–compulsive

personality disorder and perfectionism. Highly conscientious people might also be workaholics whose task orientation interferes with family and social life.

The importance of Conscientiousness for clinical psychology and psychiatry was long underrated, in part because one of the chief measures of psychopathology, the Minnesota Multiphasic Personality Inventory (MMPI), does not include clear measures of this factor. It would appear that the assessment of Conscientiousness should be of considerable value in the diagnosis of psychopathology.

Conscientiousness has implications for psychotherapy as well. Individuals who are lax and careless might not show up for therapy sessions, and they are less likely to complete homework assignments. Clinicians whose patients are low in Conscientiousness might need to make special efforts or enlist the assistance of family members to compensate for these patients' low motivation and persistence.

3.2. Physical Health

For several reasons, Conscientiousness is an important determinant of physical health. Conscientious people have the self-discipline to develop and maintain good health habits. They exercise, stick to diets, and drink in moderation. Because they are thoughtful and deliberate, they avoid unnecessary risks in driving, and they are less likely to engage in unsafe sex. Conscientious individuals are likely to be better patients who are more willing and able to comply with prescribed medical regimens. Research has generally shown higher levels of adherence among highly conscientious patients, although many other variables are also involved. For at least some of these reasons, some studies have reported an association of Conscientiousness with longevity.

Conscientiousness may sometimes be affected by physical health, particularly brain injuries or diseases. In both Alzheimer's and traumatic brain injury patients, one of the most notable features is a precipitous decline in level of Conscientiousness. Observed declines in organized and purposeful behavior may prove to be early indicators of brain disease.

3.3. Academic Performance

Research on personality and academic performance can be traced as far back as Webb's 1915 study of character and intelligence. Webb identified a factor called persistence of motives or will that clearly anticipated the construct of Conscientiousness. Many

subsequent studies confirmed that the organization, diligence, neatness, and punctuality of highly conscientious students contribute to academic success.

Academic performance is also affected by Openness to Experience, and combinations of Conscientiousness and Openness define different learning styles. Students high on both factors are prototypical good students, pursuing a wide range of interests with careful and thorough study. Students who are conscientious but closed are more likely to excel in fields such as accounting and pharmacy that emphasize accuracy over imagination.

3.4. Job Performance

Conscientiousness is best known to applied psychologists as the dimension of personality most consistently associated with job performance. A series of meta-analyses have confirmed that highly conscientious people are rated as superior performers in nearly every occupation. Other personality factors are also related to job success, but usually only in specific job categories; for example, Openness is likely to be more useful to a novelist than to a banker. But organization, persistence, meticulousness, and achievement orientation are assets in any job.

Typically, Conscientiousness scores correlate roughly .20 to .30 with performance ratings. This correlation is small, in part because performance ratings are themselves imperfect measures and in part because vocational success also depends on a host of other factors, including intelligence, training and experience, and compatibility with coworkers. Nevertheless, other things being equal, employers would be better off hiring more conscientious candidates.

If test respondents are instructed to "fake good" on a personality inventory, they typically score high in Conscientiousness. Furthermore, job applicant samples typically score higher on desirable traits than do volunteers with no incentive to manage the impressions they are making. These facts have led some industrial and organizational (I/O) psychologists to question whether self-reports on personality inventories can be trusted. However, research to date has shown that Conscientiousness scores are valid predictors even when they are administered as part of a job selection battery.

4. CONCLUSION

Conscientiousness is the dimension of personality most directly relevant to the accomplishment of tasks. Individuals who score high on this factor have clear

goals and the ability to persevere in their efforts to attain them. It is possible, of course, that one's goals are inimical to society, and highly conscientious individuals can sometimes become dangerous criminals or entrenched dictators. For the most part, however, Conscientiousness is useful to both the individual and society, and anyone interested in success in life must take it into account.

See Also the Following Articles

Intelligence Assessment ■ Intelligence, Emotional ■ Intelligence in Humans ■ Personality Assessment ■ Traits

Further Reading

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*, 261–272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9–30.

Conner, M., & Abraham, C. (2001). Conscientiousness and the theory of planned behavior: Toward a more complete model of the antecedents of intentions and behavior. *Personality and Social Psychology Bulletin*, *27*, 1547–1561.

Costa, P. T., Jr., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: Examples from Conscientiousness. *European Journal of Personality*, *12*, 117–134.

Costa, P. T., Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for Agreeableness and Conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences*, *12*, 887–898.

De Raad, B., & Schouwenburg, H. C. (Eds.). (1996). *Personality, learning, and education* [special issue]. *European Journal of Personality*, *10*(5).

Han, K., Weed, N. C., & McNeal, T. P. (1996). Searching for Conscientiousness on the MMPI-2. *Journal of Personality Assessment*, *67*, 354–363.

Ranseen, J. D., Campbell, D. A., & Baer, R. A. (1998). NEO-PI-R profiles of adults with attention deficit disorder. *Assessment*, *5*, 19–24.

Trobst, K. K., Herbst, J. H., Masters, H. L., III, & Costa, P. T., Jr. (2002). Personality pathways to unsafe sex: Personality, condom use, and HIV risk behaviors. *Journal of Research in Personality*, *36*, 117–133.

Webb, E. (1915). *Character and intelligence: An attempt at an exact study of character.* Cambridge, UK: Cambridge University Press.

Conservation Behavior

Florian G. Kaiser

Eindhoven University of Technology, Eindhoven, The Netherlands

GLOSSARY

conservation behavior Actions such as recycling, energy and water conservation, and the use of public transportation that contribute toward environmental preservation and/or conservation.

ecological footprint A measure of the totality of the environmental resources used and waste produced by a single actor (i.e., a person, a company, a nation).

face value action Behavior defined from an observer's viewpoint by its apparent consequences.

factor analysis/extraction A set of mathematical models designed to search for one or more underlying latent (i.e., unobservable) variables, the so-called factors; factors account for the intercorrelations among a set of (observed) variables.

goal-directed action Behavior defined from an actor's viewpoint as a means to achieve a subjective goal.

item A task that provokes a response that is the basis for designing variables.

item response theory (IRT) A set of measurement models in which the observed variable is understood as a function of one or more latent (i.e., unobservable) traits and one or more item features (e.g., difficulty, discrimination, guessing chance).

Rasch model The most basic model within item response theory; in the dichotomous Rasch model, the probability of an approving or correct response in an observed variable is expected to depend on two parameters only: the difficulty of the item and a latent (i.e., unobservable) trait of the respondent.

true score model The measurement model fundamental to classical test theory (CTT), the prevailing measurement model in psychology; in this model, the scores in the observed variable are a composite of the persons' (unknown) true scores and a random error.

varimax rotation A rotation method that generates orthogonal (i.e., noncorrelated) factor structures; rotation methods are designed to transform the solution from factor extraction into a solution that is easier to interpret.

Conservation behavior can be approached from either an observer's or an actor's viewpoint. When behavior is defined from the outside by its face value, even seemingly similar behaviors such as glass, paper, and battery recycling fall into distinct categories. When behavior is defined from the inside as a goal-directed performance, even diverse acts such as recycling and ownership of solar panels appear to belong to one class of actions. Correspondingly, there are two distinct research traditions that differ substantially.

1. CONCEPTIONS OF CONSERVATION BEHAVIOR

The impact of individuals' actions on the natural environment has led to ecological transformations of yet unknown gravity. Psychological analyses of

environmental problems have focused on effectively promoting more ecologically sustainable patterns and courses of action such as recycling, choice of transportation, and energy conservation. According to a recent *Journal of Social Issues* article by Stern, behaviors such as these, like other human performances, can be approached from two different angles: from an observer's viewpoint, when behavior is defined by its face value (which is evident in its consequences), and from an actor's subjective viewpoint, when behavior is defined as a motivated or goal-directed performance, namely, as the behavioral means by which people try to achieve a particular goal. Based on these two conceptions of behavior, there are two distinct research traditions that differ substantially in their behavior typology and measurement approach, in the ways in which they link behavior to its environmental consequences, and in the ways in which they tackle behavior change.

1.1. Face Value Action: A Multitude of Behavior Types

As a person drives to the grocery store, to the kindergarten, or to work, he or she is using a car that consumes fossil fuel and pollutes the air. In general, the environmental consequences—air pollution and resource consumption—are the basis for classifying a behavior as conservational or, in this case, nonconservational or environmentally harmful. An example of this view is Axelrod and Lehman's definition that categorizes "actions which contribute towards environmental preservation and/or conservation" as being conservational. Strictly speaking, however, all behaviors have environmental consequences given that every behavior uses the physical environment in which it takes place. Taking a leisurely walk down the sidewalk, for instance, involves breathing air and rubbing off microscopic pieces of concrete and rubber. Because behavior is generally consumptive, it is the relative extent and the particular quality of the environmental consumption that matter in this face value conception of conservation behavior.

When behavior is defined by its face value as it normally is, measurement is traditionally based on factor analysis, varimax rotation, and the true score (measurement) model. In this tradition, even seemingly similar behaviors such as glass, paper, and battery recycling are commonly found to fall into distinct and independent categories of conservation behavior. In turn, psychological research has focused on rather

specific behaviors, such as environmental activism, consumerism, and energy conservation, with a dominant preference for recycling behavior, as was pointed out by Vining and Ebreo. Studying behavior rather particularly, however, ultimately means that modification programs have to be specifically tailored to each type of conservation behavior.

1.2. Goal-Directed Action: One Class of Behavioral Means

Even when a behavior has environmentally harmful consequences, the individual might not necessarily intend to cause damage. At most, the individual accepts harm as a side effect of a particular performance. Thus, from a motivational viewpoint, air pollution and resource consumption often are irrelevant. For example, by taking the car to the grocery store, to the kindergarten, or to work, a person tries to increase his or her comfort level or to decrease transit time. The environmental consequences of the action are by and large ignored. However, if the person's intent is not to pollute the air or use fuel but rather to easily transport groceries home or give his or her child a safe ride, we cannot expect environmental values and attitudes to be the essential internal drives that determine car use. Instead, child-rearing and convenience-related considerations are presumably more crucial. Obviously, it is not a behavior's apparent face value that matters from a motivational viewpoint; rather, it is the personal reason, that is, the subjective goal or purpose that accompanies the act. By this logic, behavior cannot be meaningfully defined by its face value (i.e., by its obvious consequences); instead, it must be defined by its underlying motive (i.e., the goal for which a behavior is a means to an end).

When behavior is defined from a motivational viewpoint as a goal-directed performance (i.e., as the behavioral means necessary to achieve the conservation goal), research has unanimously found even seemingly diverse acts such as glass recycling and ownership of private solar panels to belong to one class of actions (based on the Rasch model within item response theory).

According to Kaiser and Wilson's model, a person's dedication to a conservation goal is most obvious in the face of increasingly demanding hurdles and progressively intolerable sacrifices. Similarly, the more obstacles a person overcomes and the more effort he or she expends along the way to the goal, the more evident

that person's commitment is to the particular cause. Likewise, if the tiniest difficulty is enough to stop a person from taking the goal-implied behavioral steps, devotion to this goal likely is rather low. In other words, the more demanding the behavioral tasks, the greater the commitment to achieving the goal implied by the performances (and vice versa).

In this second research tradition, measurement of conservation behavior essentially takes on the format of a performance test. Thus, motivationally relevant goal-directed behavior cannot be identified by the inspection of single acts. It can be measured only when several performances—all of which are necessary to achieve a particular goal—are considered and when each behavior can be characterized by the personal effort (i.e., the behavioral costs) involved in its realization. These costs are determined by the sociocultural boundary conditions (e.g., terrain, climate, customs, legal regulations, economy, societal infrastructure) in which an act takes place.

2. PERFORMANCE AND CONSEQUENCES

Regardless of whether one favors a consequences- or a goal-based conception of behavior in the conservation domain, it is the environmental impact that matters rather than conduct per se. Thus, psychology's ultimate aspiration must be to reduce people's ecological footprints in terms of pollution and energy and other resources consumed.

2.1. A Primary Focus on Consequences

In a recent *American Psychologist* article, Stern advised psychologists to justify their choice of behavior to study by its relative contribution to environmental preservation. In contrast, others decide to skip behavior entirely and to aim directly at predicting environmental consequences such as the amount of energy or water people consume. The former strategy generally diminishes psychology's significance by encouraging comparison of the impact of individual behavior with that of commercial sectors or industry. Individual households account for only approximately one-third of energy use and for only approximately 5% of solid waste. Such a focus necessarily points to the relative irrelevance of people's motives, values, and beliefs and ultimately to the relative insignificance of psychology.

Research that directly addresses the consequences rather than individual performances typically finds the boundary conditions of an act—the technology that is used, product price, and household size—to be by far the most significant determinants in affecting a person's ecological footprint. For example, regardless of any motivational differences, people consume relatively more energy for heating during the winter than during the summer. Thus, despite the fact that the absolute quantity might be decreased with an extremely conservational attitude compared with a more mundane position, the outside temperature will generally dominate the absolute amount of energy that is consumed. This relative supremacy of the external conditions is especially striking for the most environmentally crucial actions, that is, the expensive and/or high-effort performances such as retrofitting homes and purchasing energy-efficient cars.

2.2. A Deficient Relationship

Focusing on the environmental impact of people's actions obviously takes the face value definition of behavior to its extreme. However, with shared variances of 5 to 15%, the link between conservation behavior and its environmental impact is far from perfect. This finding also explains the diverse results of research based on the two discrepant criteria: behavior and its consequences. Concomitantly, if acting does not have the desired outcome, we can also conclude that people lack knowledge about the actual environmental consequences of what they themselves call conservation behavior. In other words, people may know about the general class of actions that contribute to environmental conservation and/or preservation, but at the same time they might remain ignorant of the relative environmental effectiveness of different behaviors. Logically, consequences are an unreasonable substitute for behavior.

2.3. No Goal, No Drive

Focusing too closely on the consequences ignores the actor's perspective in two essential ways. First, there are myriad environmental consequences that people should or could take into account—energy savings, biodiversity-relevant effects, greenhouse gases, and various air, soil, or water pollutants—all of which are worthwhile to consider but are far too numerous to reflect on altogether concurrently and constantly. Second, if a certain amount of energy to be saved is the ultimate goal, an actor needs to know about the

quantitative conservation effects of the various behavioral alternatives. For instance, what saves how much or relatively more energy: purchasing a fuel-efficient car or driving less? Without such knowledge, we cannot expect people to form a corresponding quantitative performance goal, and without an individual goal, there is no motive for a person to strive for it either. In other words, we cannot expect subjective reasons to be significant for a goal that does not exist. In summary, a focus on the environmental consequences makes it unlikely to find psychological determinants to be notable, and if we wish to assess psychology's factual significance in contributing to environmental conservation, psychologists must pay attention to behavior.

3. SHAPING PERFORMANCE

Despite environmentalists' and environmental scientists' arguments to the contrary, various conservation behaviors seem to fall into one mental category for a majority of people. This implies that various people who wish to act in a conservational way do not necessarily take on the same behavioral tasks to achieve their conservation goals. Whereas one person rides a bike to work, another person uses a car for daily commuting but compensates by avoiding the use of batteries. Whereas one person perceives his or her contribution to environmental preservation by switching to a vegetarian diet, another person sees some irregular recycling of bottles as his or her part. This explains why people excuse their unecological behavior by pointing to their conservational engagement. For example, an intercontinental holiday flight can subjectively be compensated by regularly buying seasonal produce, recycling, and so forth. According to Thøgersen, this interchangeability of various conservation behaviors, from a subjective viewpoint, has not received much attention yet.

3.1. Multiple Behaviors, Determinants, and Complex Interventions

When behavior is defined by its face value as it normally is, conservation behaviors are found to fall into distinct mental categories. However, if avoiding household chemicals and energy conservation are unrelated, substituting one with the other is unreasonable. Not surprisingly, various behaviors have different determinants. And predictably, an integrated theory of the

antecedents of conservation behavior has become a remote goal as well. What psychologists normally find is a multitude of influences differentially significant with various behaviors. And as it turns out, with many specific behaviors, the sociocultural boundary conditions result in the most striking effects. In terms of effect size, sociocultural boundary conditions are clearly more prominent than environmental motives, attitudes, and values, and sociocultural boundary conditions often are even more pronounced than the numerous psychological factors combined. Axiomatically, the more powerful the external conditions are, the less influential the internal ones will be.

In line with Stern in his *American Psychologist* article, conservation psychologists infer that the presence of the appropriate psychological condition, such as the right attitude, explains behavior only when there are relatively insignificant external barriers present—and even then, to a modest degree at best. Predictably, increasing sociocultural opportunities or removing such barriers to performing conservation behaviors often achieves the most powerful effects, especially when the conditions are demanding.

The psychological measures and strategies that are thought to be effective in this line of research are complex because they need to incorporate the fact that people are diverse, live under a variety of different circumstances, and respond to various interventions differentially. In contrast, intervention strategies that consist of single measures, such as a particular persuasion technique that aims at changing people's environmental attitude, often fail to have the desired behavioral consequences. However, the more sophisticated interventions become, the less efficient they are and the more time, effort, and money that are needed to develop and employ them— and, consequently, the less appealing they are for policymakers to widely promote behavior change.

3.2. Differentially Striving for a Single Goal

Because people can do different things to strive for a conservation goal, they can normally choose among various behavioral alternatives. Correspondingly, we cannot deterministically predict future behaviors; that is, we cannot foresee whether people will limit their leisure time mobility, recycle glass, or retrofit their houses. All we can safely expect is a prudent selection of the behavioral means to achieve the goal. Because any behavior being performed requires personal

resources such as time, money, and effort, we can anticipate that if two behaviors seem equally suited to achieve an objective, people will favor the relatively less demanding action over the more challenging one. Hence, if a conservation goal can be accomplished with a variety of different acts, people will most likely go with the less strenuous or less costly one.

Moreover, we can also predict that a person will engage in an act only if his or her motivation to achieve the conservation goal exceeds, or at least matches, the behavioral costs involved in its realization. In other words, the more central the conservation goal is to the person's life, the more demanding the behavioral tasks the person takes on. Predictably, Kaiser and colleagues repeatedly found that people's motivation (i.e., their intention to achieve a conservation goal) almost perfectly reflected in their goal-directed performances.

Behavior interventions, such as improving environmental motivation, generally leave people with a choice. Because the conservation goal can be accomplished with a variety of different acts, people can engage in several activities. They can buy energy-efficient light bulbs, recycle paper, refrain from using a dryer for their laundry, purchase solar panels for their own energy supply, boycott companies with poor conservation reputations, contribute to environmental organizations, and so on. Logically, because psychological measures always aim at the entire range of behavioral means, researchers cannot confirm successful behavior changes when they opt for specific and narrow performance criteria. In other words, the neglect of alternative behavioral means by which people can strive for the conservation goal makes it likely that researchers underestimate the effectiveness of psychological interventions.

Focusing on the entire range of the behavioral means implies that researchers' attention ultimately shifts away from single environmental consequences, such as the amount of heating energy or certain carbon dioxide emissions, and moves toward a person's overall environmental effect. Necessarily, with a goal-directed behavior perspective, psychology addresses the person's entire ecological footprint as well.

See Also the Following Articles

Motivation and Culture ■ Motives and Goals

Further Reading

Axelrod, L. J., & Lehman, D. R. (1993). Responding to environmental concern: What factors guide individual action? *Journal of Environmental Psychology*, *13*, 149–159.

Becker, L. J., Seligman, C., Fazio, R. H., & Darley, J. M. (1981). Relating attitudes to residential energy use. *Environment & Behavior*, *13*, 590–609.

Gatersleben, B., Steg, L., & Vlek, C. (2002). Measurement and determinants of environmentally significant consumer behavior. *Environment & Behavior*, *34*, 335–365.

Kaiser, F. G., & Wilson, M. (2004). Goal-directed conservation behavior: The specific composition of a general performance. *Personality and Individual Differences*, *36*, 1531–1544.

McKenzie-Mohr, D. (2000). Fostering sustainable behavior through community-based social marketing. *American Psychologist*, *55*, 531–537.

Stern, P. C. (2000a). Psychology and the science of human–environment interactions. *American Psychologist*, *55*, 523–530.

Stern, P. C. (2000b). Toward a coherent theory of environmentally significant behavior. *Journal of Social Issues*, *56*, 407–424.

Tanner, C., Kaiser, F. G., & Wölfing Kast, S. (2004). Contextual conditions of ecological consumerism: A food-purchasing survey. *Environment & Behavior*, *36*, 94–111.

Thøgersen, J. (1999). Spillover processes in the development of a sustainable consumption pattern. *Journal of Economic Psychology*, *20*, 53–81.

Vining, J., & Ebreo, A. (2002). Emerging theoretical and methodological perspectives on conservation behavior. In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 541–558). New York: John Wiley.

Consultation Processes in Schools

Joseph E. Zins

University of Cincinnati, Cincinnati, Ohio, USA

Charlene R. Ponti

Hamilton County Educational Services Center, Cincinnati, Ohio, USA

GLOSSARY

consultant A professional who provides consultative assistance.

consultee A person who seeks assistance from a helping professional for someone for whom they bear responsibility (e.g., a student).

ecobehavioral Theoretical orientation that combines ecological (broad viewpoint that recognizes environmental, contextual, and individual factors as influencing behavior) and behavioral (focusing on elucidating and changing inappropriate or maladaptive behavior) perspectives to help understand people's actions.

empower Acquire the means or skills to obtain resources necessary for success.

indirect service Assistance provided through a third party rather than directly as in teaching or counseling.

Consultation is the process of establishing a partnership with a parent or professional to provide problem-solving assistance regarding a child or student for whom the parent or professional is responsible. The goal is to promote students' academic and/or social-emotional well-being and performance.

1. INTRODUCTION

Most educators find that they are better able to provide the highest quality education services for their students when they work as partners with other staff and families to address student-related problems and promote competence. School consultation is a systematic process for implementing this partnership and engaging in problem solving. This article describes the knowledge, beliefs, and procedures required for successful practice.

Interest in school consultation became prominent during the late 1960s and 1970s, and there continues to be a high demand for training and practice in this area today. The impetus was that many school practitioners found that demands on their time for direct one-on-one service exceeded their capacity to provide assistance, and they also observed that many problems could be solved or prevented from becoming more serious if teachers and parents were given relatively immediate support in addressing problems that arose. Consultation addressed both of these needs, and it continues to be an important job activity for many school professionals who provide support services, including school psychologists, special educators, guidance counselors, school social workers, speech and language pathologists, and school nurses. As a result, many universities that offer professional training in education and applied psychology now offer formal coursework and practicums in this area. Further

support for such services results from the mounting body of empirical evidence demonstrating the effectiveness of consultative services.

2. DEFINITION

School consultation refers to the process of providing preventively oriented, systematic, reciprocal problem-solving assistance to empower consultees, thereby promoting students' academic and/or social–emotional well-being and performance. Participants form cooperative partnerships using an ecobehavioral perspective, with the consultees retaining responsibility for the students. The process is distinguished from "collaboration" and other services in that students are helped indirectly through the consultative problem-solving process.

3. KEY COMPONENTS AND APPLICATIONS

3.1. Participants

Consultation can involve two or more individuals. The person who provides help (e.g., school psychologist, special educator) is the consultant, and those seeking assistance are consultees (e.g., teachers, parents). The majority of school consultation takes place on a one-on-one basis, but group consultation frequently occurs as well, as with intervention assistance teams and in individualized education program (IEP) meetings or when major school reform initiatives are introduced by an external consultant. A critical distinction between consultation and other services is that the student who is the focus of the problem solving generally is not directly involved in the consultation sessions. In other words, the student is helped indirectly rather than directly as in teaching or counseling.

3.2. Ecobehavioral Perspective

Students and problems must be examined from a broad perspective because many times the issues being discussed have their origins in factors outside of the students and students live in systems outside of the school (e.g., family, community). Frequent causes of problems include the manner in which instruction is delivered, limited opportunities provided for student engagement in class, methods used for behavior management, social relationships, district policies, and family issues. However, sometimes problems are related to factors internal to students (e.g., those resulting from medical conditions such as attention deficit/hyperactivity disorder, autism, and traumatic brain injury) and to environmental factors (e.g., poverty, emotional neglect or abuse). A behavioral approach is commonly used to identify and define problems, understand the contingencies maintaining them, alter consequences to change the problem behavior, and assess the outcomes of the intervention.

3.3. Preventive Orientation

Consultation makes assistance more readily available rather than requiring a formal referral for an evaluation and then waiting for extended periods of time for it to take place as is traditionally done in many schools. Through consultation, problems can be solved in a timely manner and consultees are likely to seek assistance early before problems become more entrenched and habitual. Likewise, through the process, consultees improve their own skills and learn techniques for solving problem situations that they can apply to similar issues in the future.

3.4. Problem Solving

The essence and purpose of consultation is to solve problems. The starting point is problem identification. If problems are correctly identified, it is highly likely that they will be solved successfully. A structured systematic approach to problem solving is followed, as outlined in Table I.

3.5. Relationship and Interpersonal Aspects

Consultants exhibit warmth, genuineness, and empathy, and the consultant–consultee relationship is characterized by mutual respect, trust, and cooperation. Establishing a partnership with colleagues and parents is critical for effective consultative problem solving. What this means is that the consultant not only must view consultees as partners who possess expertise but also must actively involve and engage them as the problem-solving process unfolds. As a result, consultees are more likely to support and carry out interventions that are developed.

TABLE I
Consultative Problem-Solving Process

Establishment of cooperative partnership
- Promote understanding of each other's roles and responsibilities
- Avoid the "egalitarian virus" (failure to challenge erroneous opinions)

Problem identification and analysis
- Define problem in behavioral terms and obtain agreement with consultee
- Collect baseline data regarding problem frequency, duration, and/or intensity and conduct task analysis as needed
- Identify antecedent determinants of the problem behavior
- Identify consequences that may maintain the behavior
- Assess other relevant environmental factors
- Identify all available resources

Intervention development and selection
- Brainstorm range of possible interventions
- Evaluate the positive and negative aspects of the interventions
- Select intervention(s) from the alternatives generated

Intervention implementation, sustainability, evaluation, and follow-up
- Clarify implementation procedures and responsibilities
- Implement the chosen strategy
- Evaluate intended outcomes and any side effects
- Program generalization, plan maintenance, and develop appropriate fading procedures
- Recycle and follow up as necessary

Source. From Zins, J. E., & Erchul, W. P. (2002). Best practices in school consultation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp. 625–643). Bethesda, MD: National Association of School Psychologists. Copyright 2002 by the National Association of School Psychologists. Reprinted with permission of the publisher.

3.6. External Consultants

In addition to having special services and other staff providing consultation services, some schools hire outside experts to assist with specific projects, for example, introducing a new program such as Positive Behavioral Intervention and Support Systems and evaluating district efforts to achieve No Child Left Behind Act goal attainment. The services these individuals offer can be highly specialized, but these individuals generally work with districts on a time-limited basis.

3.7. Other Approaches to Consultation

There are many approaches to consultation in schools. Although a behavioral approach, as described in this article, is used most often in schools, mental health and organizational consultation often can be found. The essential commonalities among these approaches far exceed the differences. The term "consultation" should not, however, be confused with the "expert advice-giving" approach typically found in medical settings. The latter approach is very different from the egalitarian process described in this article.

See Also the Following Articles

Home–School Collaboration ■ School–Community Partnerships

Further Reading

Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation in applied settings.* New York: Plenum.

Caplan, G. (1970). *The theory and practice of mental health consultation.* New York: Basic Books.

Erchul, W. P., & Martens, B. K. (2002). *School consultation: Conceptual and empirical bases of practice* (2nd ed.). New York: Kluwer Academic.

Friend, M., & Cook, L. (2003). *Interactions: Collaboration skills for school professionals* (4th ed.). New York: Allyn & Bacon.

Gutkin, T. B., & Curtis, M. J. (1999). School-based consultation theory and practice: The art and science of indirect service delivery. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 598–637). New York: John Wiley.

Medway, F. J., & Updyke, J. F. (1985). Meta-analysis of consultation outcome studies. *American Journal of Community Psychology*, 13, 489–504.

Sheridan, S. M., Welch, M., & Orme, S. F. (1996). Is consultation effective? A review of outcome research. *Remedial and Special Education*, 17, 341–354.

Zins, J. E., & Erchul, W. P. (2002). Best practices in school consultation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (Vol. 1, pp. 625–643). Bethesda, MD: National Association of School Psychologists.

Zins, J. E., & Johnson, L. J. (1994). Prereferral intervention for students with special needs. In T. Husén, & T. N. Postlethwaite (Eds.), *The international encyclopedia of education, Vol. 8: Research and studies* (2nd ed., pp. 4657–4662). Oxford, UK: Elsevier Science.

Zins, J. E., Kratochwill, T. R., & Elliott, S. N. (Eds.). (1993). *Handbook of consultation services for children: Applications in educational and clinical settings.* San Francisco: Jossey–Bass.

Consumer Psychology

Michael R. Solomon

Auburn University, Auburn, Alabama, USA

GLOSSARY

archetypes Universally shared ideas and behavior patterns, involving themes such as birth, death, and the devil, that frequently appear in myths, stories, and dreams.

brand loyalty A form of repeat purchasing behavior reflecting a conscious decision to continue buying the same brand.

consumer addiction A physiological and/or psychological dependence on products or services.

consumer psychology The study of the processes involved when individuals or groups select, purchase, use, or dispose of products, services, ideas, or experiences to satisfy needs and desires.

consumer satisfaction/dissatisfaction (CS/D) The attitude a person has about a product after it has been purchased.

consumption situation Factors beyond characteristics of the product that influence a purchase decision.

diffusion of innovations The process whereby a new product, service, or idea spreads through a population.

evaluative criteria Dimensions used to judge the merits of competing options.

evoked set The options identified by a consumer that may satisfy a need.

expectancy disconfirmation model Model whereby consumers form beliefs about product performance based on prior experience with the product and/or communications about the product that imply a certain level of quality.

heuristics Mental rules-of-thumb that simplify decision making.

information search A scan of the environment to identify the options available to satisfy a need.

lifestyle A pattern of consumption reflecting a person's choices of how he or she spends time and money.

market segmentation The process of identifying groups of consumers who are similar to one another in one or more ways and devising marketing strategies that appeal to the needs of one or more of these groups.

need recognition The stage in decision making when the consumer experiences a significant difference between his or her current state of affairs and some desired state.

opinion leader A person who is frequently able to influence others' attitudes or behaviors.

perceived risk The belief that a poor product choice will produce potentially negative consequences.

psychographics Data about a person's attitudes, interests, and opinions (AIOs) that allow marketers to cluster consumers into similar groups based on lifestyles and shared personality traits.

reference group An actual or imaginary individual or group that influences an individual's evaluations, aspirations, or behavior.

self-concept The beliefs a person holds about his or her own qualities and how he or she evaluates these qualities.

self-image congruence models Models that examine the process of cognitive matching between product attributes and the consumer's self-image.

subculture A group whose members' shared beliefs and common experiences set them apart from others in the larger culture to which they belong.

symbolic self-completion theory Theory that argues that people who have an incomplete self-definition tend to bolster this identity by acquiring and displaying symbols associated with it.

values Culturally determined general ideas about good and bad goals.

Consumer psychology is the study of the processes involved when individuals or groups select, purchase, use, or dispose of products, services, ideas, or experiences to satisfy needs and desires. The decision to consume typically is the culmination of a series of stages that include need recognition, information search, evaluation of alternatives, purchase, and postpurchase evaluation. However, in some cases (especially when involvement with the product or service to be chosen is low), this rational sequence is short-circuited as consumers make decisions based on "shortcuts" called heuristics (e.g., "Choose a well-known brand name"). In other cases (especially when involvement with the product or service to be chosen is especially high, as is the case with extremely risky decisions or when the object carries extreme emotional significance to the individual), subjective criteria also may cause the person's choice to diverge from the outcome predicted by a strictly rational perspective on behavior. Indeed, many consumer behaviors, including addictions to gambling, shoplifting, and even shopping itself, are quite irrational and may literally harm the decision maker. The study of consumer psychology underscores the importance of individual and group variables that help to shape preferences for products and services. In addition to demographic differences such as age, stage in the life cycle, gender, and social class, psychographic factors such as personality traits often play a major role. A person's identification with others who constitute significant reference groups or who share the bonds of subcultural memberships also exerts a powerful impact on his or her consumption decisions. These macro influences on behavior make it more or less likely that an individual will choose to adopt new products, ideas, or services as these innovations diffuse through a market or culture.

1. I BUY, THEREFORE I AM

Consumer psychology is the study of the processes involved when individuals or groups select, purchase, use, or dispose of products, services, ideas, or experiences to satisfy needs and desires. The field embraces many kinds of consumption experiences, ranging from canned peas, a massage, or democracy to hip-hop music or a celebrity such as Madonna. Needs and desires to be satisfied range from physiological conditions such as hunger and thirst to love, status, or even spiritual fulfillment.

Consumers take many forms, ranging from an 8-year-old girl begging her mother for Pokémon cards to an executive in a large corporation deciding on a multi-million-dollar computer system. This article focuses on individual consumers, but with the caveat that many important issues relate to the psychology of group decision making involving dyads, families, and organizations. One must also recognize that the study of consumer behavior is extremely interdisciplinary. Although psychology is one of the core disciplines that have shaped the field, many other important perspectives from economics, sociology, and other social sciences also play a dominant role.

During its early stages of development, the field was often referred to as buyer behavior, reflecting an emphasis on the exchange, a transaction in which two or more organizations or people give and receive something of value. Most marketers now recognize that consumer behavior is in fact an ongoing process and not merely what happens at the moment a consumer hands over money or a credit card and, in turn, receives some good or service. This expanded view emphasizes the entire consumption process, which includes the issues that influence the consumer before, during, and after a purchase.

One of the fundamental premises of the modern field of consumer psychology is that people often buy products not for what they do but for what they mean. This principle does not imply that a product's basic function is unimportant; rather, it implies that the roles that products play in our lives extend well beyond the tasks that they perform. The deeper meanings of a product may help it to stand out from other similar goods and services. All things being equal, a person will choose the brand that has an image (or even a personality) consistent with the purchaser's underlying needs.

People's allegiances to certain sneakers, musicians, or even soft drinks help them to define their place in modern society, and these choices also enable people to form bonds with others who share similar preferences. Following are some of the types of relationships a person might have with a product:

- *Self-concept attachment.* The product helps to establish the user's identity.
- *Nostalgic attachment.* The product serves as a link with a past self.

• *Interdependence*. The product is a part of the user's daily routine.

• *Love*. The product elicits emotional bonds of warmth, passion, or other strong emotion.

Self-concept refers to the beliefs that a person holds about his or her own qualities and how he or she evaluates these qualities. People with low self-esteem may choose products that will enable them to avoid embarrassment, failure, or rejection. For example, in developing a new line of snack cakes, Sara Lee found that consumers low in self-esteem preferred portion-controlled snack items because they believed that they lacked the self-control to regulate their own eating.

A consumer exhibits attachment to an object to the extent that it is used to maintain his or her self-concept. Objects can act as a "security blanket" by reinforcing people's identities, especially in unfamiliar situations. For example, students who decorate their dorm rooms with personal items are less likely to drop out of college. This coping process may protect the self from being diluted in a strange environment. Products, especially those that serve as status symbols, also can play a pivotal role in impression management strategies as consumers attempt to influence how others think of them. Despite the adage, "You can't judge a book by its cover," in reality people often do.

The use of consumption information to define the self is especially important when an identity is yet to be adequately formed, as occurs when a person plays a new or unfamiliar role. Symbolic self-completion theory suggests that people who have an incomplete self-definition tend to bolster this identity by acquiring and displaying symbols associated with it. For example, adolescent boys may use "macho" products, such as cars and cigarettes, to augment their developing masculinity; these items act as a "social crutch" during periods of self-uncertainty.

Self-image congruence models assume a process of cognitive matching between product attributes and the consumer's self-image. Research tends to support the idea of congruence between product use and self-image. One of the earliest studies to examine this process found that car owners' ratings of themselves tended to match their perceptions of their cars; for example, Pontiac drivers saw themselves as more active and flashy than Volkswagen drivers saw themselves. Some specific attributes useful in describing matches between consumers and products include rugged/delicate, excitable/calm, rational/emotional, and formal/informal.

2. CONSUMERS AS DECISION MAKERS

Traditionally, consumer researchers have approached decision making from a rational perspective. In this view, consumers calmly and carefully integrate as much information as possible with what they already know about a product, weigh the pluses and minuses of each alternative painstakingly, and arrive at satisfactory decisions.

2.1. Stages in the Decision-Making Process

This process implies that marketing managers should carefully study the stages in decision making shown in Fig. 1 to understand how product information is obtained, how beliefs are formed, and what product choice criteria are specified by consumers. This knowledge will enable them to develop products that emphasize appropriate attributes and to tailor promotional strategies to deliver the types of information most likely to be desired in the most effective formats.

2.1.1. Need Recognition

The decision-making process begins with the stage of need recognition, when the consumer experiences a significant difference between his or her current state
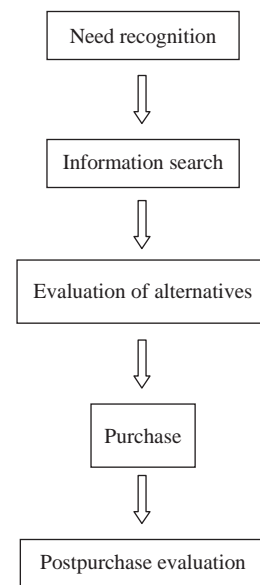

FIGURE 1 Stages in consumer decision making.

of affairs and some desired state. A person who unexpectedly runs out of gas on the highway recognizes a need, as does the person who becomes dissatisfied with the image of his or her car even though there is nothing mechanically wrong with it.

Once a need has been activated, there is a state of tension that drives the consumer to attempt to reduce or eliminate the need. This need may be utilitarian (i.e., a desire to achieve some functional or practical benefit, e.g., when a person loads up on green vegetables for nutritional reasons), or it may be hedonic (i.e., an experiential need involving emotional responses or fantasies, e.g., when a consumer thinks longingly about a juicy steak). Marketers strive to create products and services that will provide the desired benefits and permit the consumer to reduce this tension. This reduction is reinforcing, making it more likely that the consumer will seek the same path the next time the need is recognized.

Maslow's hierarchy of biogenic and psychogenic needs specifies certain levels of motives. This hierarchical approach, shown in Fig. 2, implies that one level must be attained before the next higher one is
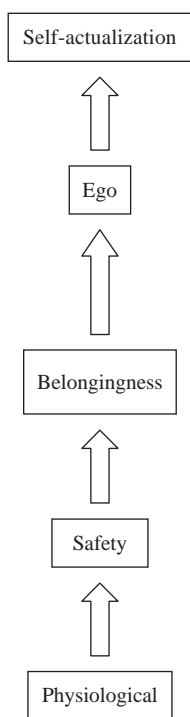
activated. Marketers have embraced this perspective because it (indirectly) specifies certain types of product benefits that people might be looking for, depending on the various stages in their development and/or their environmental conditions.

2.1.2. Information Search

Need recognition prompts information search, that is, a scan of the environment to identify the options available to satisfy the need. As a rule, purchase decisions that involve extensive search also entail perceived risk, that is, the belief that a poor choice will produce potentially negative consequences. As shown in Table I, perceived risk may be a factor if the product is expensive, complex, and hard to understand or if the consumer believes that the product will not work as promised and/or could pose a safety risk. Alternatively, perceived risk can be present when a product choice is visible to others and the consumer runs the risk of social embarrassment if the wrong choice is made.

2.1.3. Evaluation of Alternatives

Information search yields a set of alternative solutions to satisfy the need. Those identified constitute the consumer's evoked set. How does a consumer decide which criteria are important, and how does he or she narrow down product alternatives to an acceptable number and eventually choose one instead of others? The answer varies depending on the decision-making process used. A consumer engaged in extended problem solving may carefully evaluate several brands, whereas someone making a habitual decision might not consider any

FIGURE 2 Maslow's hierarchy of needs. Adapted from Maslow, A. H. (1970). *Motivation and personality* (2nd ed.). New York: Harper & Row.

TABLE I
Types of Perceived Risk

Monetary risk	Expensive products
Functional risk	Products requiring time or effort to understand
Physical risk	Mechanical products, medical products, food
Social risk	Socially visible products or those that can influence the impressions made on others (e.g., deodorant)
Psychological risk	Expensive luxuries that may involve guilt or addiction

Source. Solomon, M. R. (2004). *Consumer behavior: Buying, having, and being* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

alternatives to his or her normal brand. Variety seeking, or the desire to choose new alternatives over more familiar ones, can also play a role; consumers at times are willing to trade enjoyment for variety because the unpredictability itself is rewarding.

Evaluative criteria are the dimensions used to judge the merits of competing options. If all brands being considered rate equally well on one attribute (e.g., if all televisions come with remote control), consumers will have to find other reasons to choose one over the others. Determinant attributes are the characteristics actually used to differentiate among choices. For example, consumer research by Church & Dwight Company indicated that many consumers view the use of natural ingredients as a determinant attribute when selecting personal care products. This prompted the firm to develop toothpaste made from baking soda, an ingredient that the company already manufactured for its Arm & Hammer brand.

2.1.4. Purchase

Once a need has been recognized, a set of feasible options (often competing brands) that will satisfy the need have been identified, and each of these options has been evaluated, the "moment of truth" arrives: The consumer must make a choice and actually procure the product or service. However, other factors at the time of purchase may influence this decision. A consumption situation is defined by factors beyond characteristics of the product that influence a purchase decision. These factors can be behavioral (e.g., entertaining friends) or perceptual (e.g., being depressed, feeling pressed for time).

A consumer's mood can have a big impact on purchase decisions. For example, stress can impair information-processing and problem-solving abilities. The two dimensions of pleasure and arousal determine whether a shopper will react positively or negatively to a consumption situation. In addition, the act of shopping itself often produces psychological outcomes ranging from frustration to gratification or even exhilaration.

Despite all of their efforts to "pre-sell" consumers through advertising, marketers increasingly recognize that many purchases are strongly influenced by the purchasing environment. Indeed, researchers estimate that shoppers decide on approximately two of every three products while wheeling their carts through supermarket aisles. Dimensions of the physical environment, such as decor, ambient sounds or music, and even temperature, can influence consumption significantly. One study even found that pumping in certain odors in a Las Vegas casino actually increased the amount of money that patrons fed into slot machines. Time is another important situational variable. Common sense dictates that more careful information search and deliberation occurs when consumers have the luxury of taking their time.

2.1.5. Postpurchase Evaluation and Satisfaction

Consumer satisfaction/dissatisfaction (CS/D) refers to the attitude that a person has about a product after it has been purchased. This attitude, in turn, is an important determinant of whether the item will be bought again in the future. Despite evidence that customer satisfaction is steadily declining in many industries, marketers are constantly on the lookout for sources of dissatisfaction. For example, United Airlines' advertising agency set out to identify specific aspects of air travel that contributed to discontent during the travel experience. The agency gave frequent fliers crayons and a map showing different stages in a long-distance trip and asked passengers to fill in colors using hot hues to symbolize areas causing stress and anger and using cool colors for parts of the trip associated with satisfaction and calm feelings. Although jet cabins tended to be filled in with a serene aqua color, ticket counters were colored orange and terminal waiting areas were colored fire red. This research led the airline to focus more on overall operations instead of just on in-flight experiences, and the "United Rising" advertising campaign was born.

Satisfaction is not determined solely by the actual performance quality of a product or service. It is also influenced by prior expectations regarding the level of quality. According to the expectancy disconfirmation model, consumers form beliefs about product performance based on prior experience with the product and/or communications about the product that imply a certain level of quality. When something performs the way in which consumers thought it would, they might not think much about it. If, on the other hand, the product fails to live up to expectations (even if those expectations are unrealistic), negative affect may result. If performance happens to exceed their expectations, consumers are satisfied and pleased. This explains why companies sometimes try to "under-promise" what they can actually deliver.

2.2. Biases in the Decision-Making Process

Although the rational model of decision making is compelling, many researchers now recognize that decision makers actually possess a repertoire of strategies—and not all of these strategies are necessarily rational. The constructive processing perspective argues that a consumer evaluates the effort required to make a particular choice and then chooses a strategy best suited to the level of effort required.

As shown in Table II, some purchases are made under conditions of low involvement, where the consumer is not willing to invest a lot of cognitive effort. Instead, the consumer's decision is a learned response to environmental cues, for example, when he or she impulsively decides to buy something that is promoted as a "surprise special" in a store. In other cases, the consumer is highly involved in a decision, and again the stages of rational information processing might not capture the process. For example, the traditional approach is hard-pressed to explain a person's choice of art, music, or even a spouse. In these cases, no single quality may be the determining factor. Instead, an experiential perspective stresses the Gestalt, or totality, of the product or service.

TABLE II
Limited Versus Extended Problem Solving

Motivation	Low risk	High risk
Information search	Low search	High search
Evaluation of alternatives	Weak beliefs	Strong beliefs
	Few differences perceived among alternatives	Significant differences perceived among alternatives
	Most prominent criteria or heuristics used	Many criteria used
Purchase	Limited time spent	Extensive time spent
	Few stores shopped	Many stores shopped
	Store displays influential	Advertising and store personnel influential

Source. Solomon, M. R. (2004). *Consumer behavior: Buying, having, and being* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Consumption at the low end of involvement typically is characterized by inertia, where decisions are made out of habit because the consumer lacks the motivation to consider alternatives. Many people tend to buy the same brand nearly every time they go to the store. A competitor who is trying to change a buying pattern based on inertia often can do so rather easily because little resistance to brand switching will be encountered if the right incentive is offered.

At the high end of involvement, one can expect to find the type of passionate intensity that is reserved for people and objects that carry great meaning for the individual. When consumers are truly involved with a product, an ad, or a Web site, they enter a flow state. Flow is an optimal experience characterized by a sense of playfulness, a feeling of being in control, highly focused attention, and a distorted sense of time.

Especially when limited problem solving occurs prior to making a choice, consumers often fall back on heuristics, that is, mental rules-of-thumb that lead to a speedy decision. These rules range from the very general (e.g., "Higher priced products are higher quality products," "Buy the same brand I bought last time") to the very specific (e.g., "Buy Domino, the brand of sugar my mother always bought").

One frequently used shortcut is the tendency to infer hidden dimensions of products from observable attributes. These are known as product signals. Country of origin is an example of a commonly used product signal. In some cases, people may assume that a product made overseas is of better quality (e.g., cameras, cars), whereas in other cases, the knowledge that a product has been imported tends to lower perceptions of product quality (e.g., apparel). Price is also a heuristic; all things equal, people often assume that "Brand A" is of higher quality simply because it costs more than "Brand B."

A well-known brand also frequently functions as a heuristic. People form preferences for a favorite brand and then literally might never change their minds in the course of a lifetime. In contrast to inertia, brand loyalty is a form of repeat purchasing behavior reflecting a conscious decision to continue buying the same brand. Purchase decisions based on brand loyalty also become habitual over time, although in these cases the underlying commitment to the product is much more firm. Because of the emotional bonds that can come about between brand-loyal consumers and products, "true blue" users react more vehemently when these products are altered, redesigned, or eliminated. For example, when Coca-Cola replaced its tried-and-true

formula with New Coke during the 1980s, the company encountered a firestorm of national call-in campaigns, boycotts, and other protests.

Finally, many of people's reactions to products are based on aesthetic responses to colors, shapes, and objects. Many of these preferences are deep-seated or culturally determined. Package designs often incorporate extensive research regarding consumers' interpretations of the meanings accorded to symbols on the box or can. These meanings may be subtle or blatant, but they can exert a powerful effect on expectations about the product within the box or can. In one study, respondents rated the taste of a beer as heavier and more robust when it was served in a brown glass bottle than when the same product was dispensed in a clear bottle.

2.3. "Irrational" Decision Making

Psychologists have identified many cognitive mechanisms that interfere with "objective" information processing and decision making. Other researchers have gone a step further in their focus on domains of consumer behavior that cannot be readily explained by a cognitive perspective or where an individual's actions are actually irrational and perhaps even dysfunctional. Indeed, some of consumers' buying behaviors do not seem rational at all because they do not serve a logical purpose (e.g., collectors who pay large sums of money for paraphernalia formerly owned by rock stars). Other actions occur with virtually no advance planning at all (e.g., impulsively grabbing a tempting candy bar from the rack while waiting to pay for groceries).

Still other consumer behaviors, such as excessive eating, excessive drinking, and cigarette smoking, are actually harmful to the individual. These actions may be facilitated by many psychological factors, including the desire to conform to the expectations of others, learned responses to environmental cues, and observational learning prompted by exposure to media. The cultural emphasis on wealth as an indicator of self-worth may encourage activities such as shoplifting and insurance fraud. Exposure to unattainable ideals of beauty and success in advertising can create dissatisfaction with the self, sometimes resulting in eating disorders or self-mutilation.

Consumer addiction is a physiological and/or psychological dependence on products or services. Although most people equate addiction with drugs, virtually any product or service can be the focus of psychological dependence; for example, there is even a support group for Chap Stick addicts. Some psychologists are now voicing strong concerns about "Internet addiction," a condition whereby Web surfers become obsessed by online chat rooms to the point where their "virtual" lives take priority over their real ones. Compulsive consumption refers to repetitive excessive shopping that serves as an antidote to tension, anxiety, depression, or boredom.

3. PSYCHOLOGICAL INFLUENCES ON CONSUMPTION

The business process of market segmentation identifies groups of consumers who are similar to one another in one or more ways and then devises marketing strategies that appeal to the needs of one or more of these groups. One very common way in which to segment consumers is along demographic dimensions such as the following:

- *Age.* People who belong to the same age group tend to share a set of values and common cultural experiences that they carry throughout life.
- *Gender.* Many products are sex typed, and consumers often associate them with one gender or the other. Marketers typically develop a product to appeal to one gender or the other. For example, Crest's Rejuvenating Effects toothpaste, made specifically for women, is packaged in a teal tube nestled inside a glimmering "pearlescent" box.
- *Social class and income.* Working-class consumers tend to evaluate products in more utilitarian terms such as sturdiness and comfort. They are less likely to experiment with new products or styles such as modern furniture and colored appliances. Higher classes tend to focus on more long-term goals such as saving for college tuition and retirement.
- *Family structure.* People's family and marital status influences their spending priorities. For example, young bachelors and newlyweds are the most likely to exercise, consume alcohol, and go to bars, concerts, and movies.
- *Race and ethnicity.* Ethnic minorities in the United States spend more than $600 billion per year on products and services, so firms must devise products and communications strategies tailored to the needs of these subcultures. Sometimes these differences are subtle yet important. When Coffee-Mate discovered that African Americans tend to drink their coffee with sugar and cream more so than do Caucasians, the company mounted a promotional blitz in black-oriented media that resulted in double-digit increases in sales volume.

• *Geography*. Place of residence influences preferences within many product categories, from entertainment to favorite cars, decorating styles, or leisure activities. For example, BMW found that drivers in France prized its cars for their road-handling abilities and the self-confidence this gave them, whereas drivers in Austria were more interested in the status aspect of the BMW brand.

3.1. Psychographics and Lifestyles

Although these segmentation variables are very important, consumers can share the same demographic characteristics and still be very different people. Psychographics are data about people's attitudes, interests, and opinions (AIOs) that allow marketers to cluster consumers into similar groups based on lifestyles and shared personality traits.

Lifestyle refers to a pattern of consumption reflecting a person's choices of how he or she spends time and money. In an economic sense, a person's lifestyle represents the way in which he or she has elected to allocate income in terms of relative allocations to various products and services and to specific alternatives within these categories. Lifestyle, however, is more than the allocation of discretionary income; it is a statement about who a person is in society and who the person is not. Group identities, whether of hobbyists, athletes, or drug users, gel around forms of expressive symbolism.

3.2. Personality Theory Applications to Consumer Behavior

A consumer's personality, or unique psychological makeup, may influence the products and marketing messages that he or she prefers. Consumer psychologists have adapted insights from major personality theorists to explain people's consumption choices.

3.2.1. Freudian Theory

Freudian psychology exerted a significant influence on applied consumer research, especially during the early days of the discipline. Freud's writings highlight the potential importance of unconscious motives underlying purchases. This perspective also hints at the possibility that the ego relies on the symbolism in products to compromise between the demands of the id and the prohibitions of the superego. During the 1950s, motivational researchers attempted to apply Freudian ideas

to understand the deeper meanings of products and advertisements. For example, for many years, Esso (now Exxon) reminded consumers to "Put a Tiger in Your Tank" after researchers found that people responded well to this powerful animal symbolism containing vaguely sexual undertones.

3.2.2. Jungian Theory

Freud's disciple, Jung, introduced the concept of the collective unconscious, that is, a storehouse of memories inherited from a person's ancestral past. These shared memories create archetypes, that is, universally shared ideas and behavior patterns involving themes such as birth, death, and the devil, that frequently appear in myths, stories, and dreams. For example, some of the archetypes identified by Jung and his followers include the "old wise man" and the "earth mother," and these images appear frequently in marketing messages that use characters such as wizards, revered teachers, and even Mother Nature.

3.2.3. Trait Theory

Trait theory focuses on the quantitative measurement of personality traits, that is, identifiable characteristics that define a person. Some specific traits relevant to consumer behavior include innovativeness (i.e., the degree to which a person likes to try new things), materialism (i.e., the amount of emphasis placed on acquiring and owning products), self-consciousness (i.e., the degree to which a person deliberately monitors and controls the image of the self that is projected to others), need for cognition (i.e., the degree to which a person likes to think about things and, by extension, expend the necessary effort to process brand information), and self-monitoring (i.e., the degree to which a person is concerned with the impression that his or her behaviors make on others).

4. CULTURAL AND INTERPERSONAL INFLUENCES ON CONSUMPTION

Values are very general ideas about good and bad goals, and these priorities typically are culturally determined. From these flow norms, that is, rules dictating what is right or wrong and what is acceptable or unacceptable. Consumers purchase many products and services because they believe that these products will help them

to attain value-related goals. For example, an emphasis on personal hygiene in Japan has created a demand for products such as automated teller machines (ATMs) that literally "launder" money by sanitizing yen before dispensing them to bank customers.

4.1. Subcultures and Reference Groups

Members of a subculture share beliefs and common experiences that set them apart from others in the larger culture. Whether "Dead Heads" or "Skinheads," each group exhibits its own unique set of norms, vocabulary, and product insignias (e.g., the skull and roses that signifies the Grateful Dead subculture).

A reference group is an actual or imaginary individual or group that influences an individual's evaluations, aspirations, or behavior. As a rule, reference group effects are more robust for purchases that are (a) luxuries (e.g., sailboats) rather than necessities and (b) socially conspicuous or visible to others (e.g., living room furniture, clothing).

4.2. Opinion Leaders

An opinion leader is a person who is frequently able to influence others' attitudes or behaviors. Opinion leaders are extremely valuable information sources because (a) they are technically competent and, thus, more credible; (b) they have prescreened, evaluated, and synthesized product information in an unbiased way; and (c) they tend to be socially active and highly interconnected in their communities.

4.3. Diffusion of Innovations

Diffusion of innovations refers to the process whereby a new product, service, or idea spreads through a population. If an innovation is successful (most are not), it typically is initially bought or used by only a few people. Then, more and more consumers decide to adopt it until it might seem (sometimes) that nearly everyone has bought or tried the innovation.

A consumer's adoption of an innovation resembles the standard decision-making sequence whereby he or she moves through the stages of awareness, information search, evaluation, trial, and adoption. The relative importance of each stage may differ depending on how much is already known about the innovation as well as on cultural factors that may affect people's willingness to try new things. However, even within the same culture, not all people adopt an innovation at the same rate. Some do so quite rapidly, whereas others never do at all.

Consumers can be placed into approximate categories based on their likelihood of adopting an innovation. Roughly one-sixth of the people are very quick to adopt new products (i.e., innovators and early adopters), and one-sixth of the people are very slow (i.e., laggards). The other two-thirds are somewhere in the middle (i.e., late adopters). These latter consumers are the mainstream public; they are interested in new things, but they do not want them to be too new.

Even though innovators represent only approximately 2.5% of the population, marketers are always interested in identifying them. Innovators are the brave souls who are always on the lookout for novel developments and who will be the first to try new offerings. They tend to have more favorable attitudes toward taking risks, have higher educational and income levels, and be socially active. In addition, many innovators also are opinion leaders, so their acceptance of an innovation may be a crucial factor in persuading others to try it as well.

As a rule, consumers are less likely to adapt innovations that demand radical behavior changes—unless they are convinced that the effort will be worthwhile. As a result, evolutionary changes (e.g., a cinnamon version of Quaker oatmeal) are more likely to be rapidly adapted than are revolutionary changes (e.g., ready-to-eat microwaveable Quaker oatmeal). The following factors make it more likely that consumers will accept an innovation:

- Compatibility with current lifestyle
- Ability to try the product before buying
- Simplicity of use
- Ease of observing others using the innovation
- Relative advantage over benefits offered by other alternatives

See Also the Following Articles

Advertising and Culture ■ Advertising Psychology ■ Economic Behavior ■ Values and Culture

Further Reading

Alba, J. W., & Hutchinson, J. W. (1988). Dimensions of consumer expertise. *Journal of Consumer Research*, *13*, 411–454.

Baumgartner, H. (2002). Toward a personology of the consumer. *Journal of Consumer Research*, *29*, 286–292.

Belk, R. W. (1988). Possessions and the extended self. *Journal of Consumer Research*, *15*, 139–168.

Belk, R. W., Wallendorf, M., & Sherry, J. F., Jr. (1989). The sacred and the profane in consumer behavior: Theodicy on the odyssey. *Journal of Consumer Research*, *16*, 1–38.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1988). Constructive consumer choice processes. *Journal of Consumer Research*, *25*, 187–217.

Bettman, J. R., & Park, C. W. (1980). Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *Journal of Consumer Research*, *7*, 234–248.

Cialdini, R. B. (1993). *Influence: The psychology of persuasion*. New York: William Morrow.

Dichter, E. (1964). *The handbook of consumer motivations*. New York: McGraw–Hill.

Fournier, S. (1998). Consumers and their brands: Developing relationship theory in consumer research. *Journal of Consumer Research*, *24*, 343–373.

Hoffman, D. L., & Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing*, *60*, 50–68.

Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage.

Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, *9*, 132–140.

Jacoby, J., & Chestnut, R. (1978). *Brand loyalty: Measurement and management*. New York: John Wiley.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Maslow, A. H. (1970). *Motivation and personality* (2nd ed.). New York: Harper & Row.

Midgley, D. F. (1983). Patterns of interpersonal information seeking for the purchase of a symbolic product. *Journal of Marketing Research*, *20*, 74–83.

Oliver, R. L. (1996). *Satisfaction*. New York: McGraw–Hill.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*, 135–146.

Ratner, R. K., Kahn, B. E., & Kahneman, D. (1999). Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research*, *26*, 1–15.

Robertson, T. S., & Kassarjian, H. H. (Eds.). (1991). *Handbook of consumer behavior*. Englewood Cliffs, NJ: Prentice Hall.

Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed.). New York: Free Press.

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Sirgy, M. J. (1982). Self-concept in consumer behavior: A critical review. *Journal of Consumer Research*, *9*, 287–300.

Solomon, M. R. (1983). The role of products as social stimuli: A symbolic interactionism perspective. *Journal of Consumer Research*, *10*, 319–329.

Solomon, M. R. (2004). *Consumer behavior: Buying, having, and being* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Snyder, M. (1979). Self-monitoring processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 85–128). New York: Academic Press.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, *4*, 199–214.

Wells, W. D., & Tigert, D. J. (1971). Activities, interests, and opinions. *Journal of Advertising Research*, *11*, 27.

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion*. Hillsdale, NJ: Lawrence Erlbaum.

Wright, P. (2002). Marketplace metacognition and social intelligence. *Journal of Consumer Research*, *28*, 677–682.

Zablocki, B. D., & Kanter, R. M. (1976). The differentiation of life-styles. *Annual Review of Sociology*, *2*, 269–297.

Contracts

René Schalk

Tilburg University, Tilburg, The Netherlands

GLOSSARY

obligation A commitment to some future action.
relational relationship Focusing on open-ended relationships involving considerable investments of all parties.
transactional relationship Focusing on short-term and monetizable exchanges.
violation A failure to comply with the terms of the contract.
zone of negotiability Room at the level of the individual party for negotiation about contract terms.

A contract is an exchange agreement between parties. In employment, psychological contracts are drivers of employee behavior with a strong influence on organizational outcomes.

1. INTRODUCTION

Contracts are the cornerstones of our society. Contracts about exchanges are a common feature of societies all over the world. We seem not to be able to do without contracts. In our private lives, we have marriage contracts and other forms of contractual arrangements that people make when they decide to live together. In our social lives, we have contracts in the form of sports club memberships, political party memberships, and the like. We make contractual exchanges when buying houses, booking holidays, paying for insurance, and so on. At work, we have labor contracts. Contracts play a very prominent role in our daily lives. One could say that, in a way, we are nearly permanently occupied with fulfilling contract terms. Therefore, it is inevitable that we will sometimes be confronted with unfulfilled or broken contracts by ourselves or by other parties.

2. WHAT IS A CONTRACT?

The belief of obligations existing between two or more parties is the basis for a contract. A contract is an exchange agreement between two (or more) parties. The contract may include written terms, orally communicated terms, or other expressions of commitment and future intent. As generally defined, a contract is based on a mutually agreed on, voluntarily made exchange of promises between two (or more) parties, where each party gains if the exchange agreement is fulfilled and each party loses if it is not.

3. DIFFERENT TYPES OF CONTRACTS

Contracts can be viewed from various angles. The economic focus is on transaction or relationship costs of the contractual exchange relationships. The legal focus is on the formalized terms and procedures and on issues related to contract fulfillment. The term "social contract" refers to a pact or an agreement made by all of the individuals who are to compose it.

The term "psychological contract" refers to ideas about mutual obligations in the exchange agreement. One view on the psychological contract is that it includes the perceptions of both parties. Another view limits the psychological contract to individual beliefs about the exchange agreement. It includes and adds to the economic and legal perspectives. The social contract is based on generalized psychological contracts.

Psychological contracts address the important issue of incompleteness of contracts. It is virtually impossible to create complete formal contracts. Unforeseen events may arise over time. Foreseen events cannot always be described unambiguously, and the costs of negotiating all possible events and consequences might be too high. Psychological contracts are generally understood agreements that fill in the gaps of the formal agreement.

A general distinction can be made between transactional contracts, which have the characteristics of an "arm's length" market transaction (e.g., between buyer and seller), and relational contracts such as complex long-term relationships between employer and employee.

4. CONTRACTS IN EMPLOYMENT

A basic feature of employee–employer interaction is the nature of the exchange in the employment relationship. An employment relationship is a contract based on a mutually agreed on, voluntarily made exchange of promises between two (or more) parties, where each party gains if the exchange agreement is fulfilled and each party loses if it is not. Each component can vary in response to the social context in which employment is embedded.

A meeting of the minds of employer and employee is implied in the term "mutuality." Whether there is actually a moment when parties share a crystal clear understanding regarding the terms of their exchange is open to question. Nevertheless, at the start of the employment relationship, there is at least a perception of agreement if not necessarily agreement in fact.

Both employer and employee have free choice in entering into and accepting the terms of an employment arrangement. However, this is always a matter of degree. Voluntariness in employment is historically rooted in the evolution of institutions such as individual rights, ownership of private property, and collective bargaining. Voluntariness in employment is possible when individuals have the right to control their own time, services, and use of their skills. Employer beliefs regarding worker freedom of choice (e.g., job opportunities elsewhere) influence the inducements that employers offer in exchange for worker contributions.

5. CONSEQUENCES OF VIOLATIONS OF CONTRACTS

Fulfilling contract terms by the organization leads to greater employee commitment to the organization and to the relationship. Consequently, readiness to accept work roles and tasks, willingness to engage in extra-role behaviors, and willingness to avoid negative behaviors (e.g., coming to work late, not performing well) will be greater. This will result in more committed, motivated, and trusting organizational citizens.

More favorable psychological contracts are related to a higher degree of organizational commitment and a lower degree of intention to leave the organization. On the other hand, when employees believe that the organization has not fulfilled its side of the bargain, they are less likely to engage in extra-role behaviors (e.g., organizational citizenship). Employees who believe that the organization has violated certain terms of the psychological contract will be more likely to have high rates of tardiness and absenteeism and a higher intention to leave the organization. Anger, disappointment, and strong emotions may also be consequences of not fulfilling contract terms.

6. CROSS-NATIONAL DIFFERENCES

Who the parties to the employment agreement are varies considerably across countries, as does the actual level at which the exchange agreement is created. Societies ascribe varying importance to the parties of an employment contract. The status and esteem of government and the state per se differ considerably. Because parties to employment agreements entail a wide range of societal actors, employment agreements can arise at several

levels: between the individual worker and the firm, between groups of workers and the firm, between groups of workers and groups of firms, and in central agreements involving groups of workers, employers, and the state. An industrial relations system involves three sets of actors: employers, workers (typically in groups either within or between firms), and the government (including direct negotiation involving governmental officials, governmental mediation of employee–employer agreements, or creation of laws and statutes specifying conditions of employment).

Three general types of employment and industrial relations systems are direct exchanges between firm and employee; central agreements between employer organizations and unions (or comparable parties), with possible segregation at various levels (e.g., nationwide, type of industry); and employment relationships anchored in society (i.e., the state and institutions), where regulations and statutes predominate in the construction of the employee–employer exchange.

Direct firm–employee exchanges are more prevalent in countries such as Israel, New Zealand, Singapore, and the United States (except in the case of unionized firms). Central agreements play an important role in countries such as Australia, Belgium, India, Mexico, The Netherlands, and Sweden. Regulations play a significant role in countries such as France and Japan.

One striking difference across societies is in the zone of negotiability of the individual employment relationship. Factors shaping it include the individual's personal power in constructing an idiosyncratic arrangement with the employer, the employer's willingness to demand or offer unique or varied employment conditions to individual workers, and the society's regulation of employment and tolerance for unequal outcomes. State intervention or central agreements have the most salient effect on the range and boundaries for what is negotiable between employers and workers. State-level interventions have several effects. Regulations constrain what conditions are negotiable and mandate others. Central agreements, negotiated collectively, limit individual-level variability in contract terms. Both appear to be more prevalent in societies that are relatively intolerant of outcome inequality or where there are institutional concerns regarding disparities in power between labor, on the one hand, and management and owners, on the other. Laws stipulate which conditions are non-negotiable, for example, requiring payment of retirement or unemployment taxes and providing extended family leave benefits following childbirth. Laws also serve to enforce negotiated terms of employment.

Nonlegal factors can also constrain the zone of negotiability. Cultural norms, in particular, can effectively remove certain employment terms from the bargaining table. In societies with strong state influence or central agreements, the focus of negotiation and flexibility shifts to what is possible within or around the rules. There tends to be a relatively clear demarcation of which issues are on the negotiation table and which ones are not. Yet even when the wage system has been standardized with strict rules for salary ranges for certain jobs, there is often room at the level of the individual worker for negotiation about salary within the given range or about the title of a particular job.

7. IMPLICATIONS

Contractual agreements play a very prominent role in our society and in the world of work. Contractual exchange agreements at work determine employee behavior and organizational outcomes.

In the contracting process clear communication and careful assessment of the contract terms is necessary; constant monitoring and renegotiating, adjusting to changing needs of the organization and the individual employee.

In the future, we can expect considerable innovation and experimentation in contracting. Critical questions will center on how to balance differentiation and fairness, especially in situations where employees are interdependent, and how to integrate groups and promote cooperation while keeping clear boundaries for individual psychological contracts. A greater variety of contract types, including more "individualized" results of contract negotiations, will exist. Contracts made today are a way in which to both know and shape the future. The management of psychological contracts is a core task for firms that attempt to develop "people-building" rather than "people-using" organizations in an organizational climate characterized by trust. Increasingly, managing the psychological contract is a core task for workers themselves, who seek to meet their own needs in active individual- and group-level negotiations with their employers.

See Also the Following Articles

Job Analysis, Design, and Evaluation ■ Occupational Choice ■ Occupational Psychology, Overview ■ Part-Time Work

Further Reading

Anderson, N., & Schalk, R. (1998). The psychological contract in retrospect and prospect. *Journal of Organizational Behavior, 19,* 637–648.

Herriot, P., & Pemberton, C. (1995). *New deals: The revolution in managerial careers.* Chichester, UK: Wiley.

Rousseau, D. M. (1995). *Psychological contracts in organizations: Understanding written and unwritten agreements.* Thousand Oaks, CA: Sage.

Rousseau, D. M., & Schalk, R. (Eds.). (2000). *Psychological contracts in employment: Cross-national perspectives.* Thousand Oaks, CA: Sage.

Schalk, R., & Rousseau, D. M. (2001). Psychological contracts in employment. In N. Anderson, D. S. Ones, H. Kepir Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology, Vol. 2: Organizational psychology.* Thousand Oaks, CA: Sage.

Cooperation at Work

Eduardo Salas, Dana E. Sims, and Cameron Klein

University of Central Florida, Orlando, Florida, USA

GLOSSARY

adaptability The ability to gather information from the task environment and use it to adjust the team's strategy through the use of compensatory behavior and the reallocation of interteam resources.

cooperation Includes offering help to only those team members who need it, pacing activities to fit the needs of the team, and behaving in an unambiguous manner so that actions are not misinterpreted.

cross-training A training strategy in which teammates develop an understanding for the tasks, mission, objectives, and responsibilities of the team; this strategy targets team members' interpositional knowledge and shared mental models for development with resulting gains in team coordination.

mutual performance monitoring Monitoring fellow team members' performance while maintaining individual responsibilities to improve overall group performance.

shared mental models (SMMs) Shared understandings or representations of the goals of the team, individual team member tasks, and how team members will coordinate to achieve their common goals; individual team members can have a varying degree of overlap or "sharedness" among their mental model of the team.

team Collective of interdependent individuals who have shared habits of thinking, mental models, and paradigms that guide their perceptions and behaviors toward a common goal.

team building Targets role clarification, goal setting, problem solving, and/or interpersonal relations for improvement.

team coordination training Focuses on teaching team members about basic processes underlying teamwork; this strategy targets mutual performance monitoring and back-up behavior as well as other critical teamwork competencies and skills for improvement.

team orientation The preference for working collectively and the tendency to enhance individual performance by using task inputs from other group members during group tasks.

team self-correction training Team members are taught techniques for monitoring and then categorizing their own behaviors; this process generates instructive feedback so that team members can review performance episodes and correct deficiencies.

team structure Refers to the power structure (e.g., hierarchical), the number of team members, and the diversity among them.

Teamwork is a set of flexible behaviors, cognitions, and attitudes that interact to achieve desired mutual goals and adaptation to changing internal and external environments. It consists of the knowledge, skills, and attitudes (KSAs) that are displayed in support of one's teammates and the team's objectives. Essentially, teamwork is a set of interrelated thoughts, actions, and feelings that combine to facilitate coordinated adaptive

performance and the completion of taskwork objectives. When teamwork KSAs are mastered and displayed, teams can achieve a level of performance and effectiveness that is greater than the total efforts of all individual team members.

1. IMPORTANCE OF TEAMWORK

The nature of work is changing. Recent technological advances, a shift from manufacturing to service-based organizations, increased global competition, and the importance of knowledge workers all have contributed to a dynamic and complex work environment. To survive, organizations must embrace flexibility and adaptability. They must have a systems view that helps them to integrate the meaning of common everyday occurrences into overall interrelationships, interdependencies, and patterns of change to allow them to achieve this needed adaptability. Very often, organizations accomplish these goals through the implementation of teams. It is in direct response to increasingly severe time pressures, technological advancements, and other elements of highly complex environments that teamwork has become an established stronghold and a source of competitive advantage for today's organizations. Teams are also called on when the task demands are such that the capabilities of one individual are exceeded or the consequences of error are high. Therefore, the utility of teams and teamwork has spanned industry to include aviation, nuclear power, and the military.

A team is a collective of interdependent individuals who together have shared objectives, mental models, and procedures that guide their perceptions, thinking, and behaviors toward a common goal. The process in which teams achieve this commonality is called collaboration. Similarly, cooperation has been discussed as a team skill competency that includes offering help to those team members who need it, pacing activities to fit the needs of the team, and behaving so that actions are not misinterpreted. Collective organizations immerse themselves in the virtues of collaboration and cooperation by reducing the hierarchy of social stratification on which traditional bureaucratic organizations rely. Furthermore, they encourage decisions to be made by the collective group to foster a sense of community and shared purpose. Cooperation and teamwork facilitate coordination, communication, adaptability, enhanced employee participation, and empowerment, thereby allowing individuals to achieve collective outputs that are greater than the sum of their parts.

Organizations continue to depend on teams and the synergy they engender to assist in streamlining work processes that promote efficiency, increased innovation, and quality products and services.

This article outlines the nature of cooperation and teamwork as it is manifested in organizations. It examines what derails teamwork and what facilitates it as well as how to promote it through training. The article concludes with a discussion of the need to measure training effectiveness and offers a brief discussion on how this is done. Because teams and teamwork will likely continue to play a vital role in organizations into the future, it is a worthy endeavor to examine their nature.

2. TEAMWORK EXAMINED

Despite extensive writing on the topic, a clear definition and description of teamwork continues to be elusive. In general, teamwork is more than just the summated task or task-related accomplishments of the individuals who make up a group. It is the interaction of team-related knowledge, skills, and attitudes (KSAs) of all team members—and the ability to capitalize on this pool of resources—that maximizes the potential effectiveness of the group to outperform even the most competent individual member of the team. The definition of teamwork provided here is a synthesis of previous definitions that have been advanced by established team researchers. Therefore, teamwork is defined as a set of flexible behaviors, cognitions, and attitudes that interact to facilitate taskwork and achieve mutually desired goals and adaptation to the changing internal and external environments. Although many behaviors, cognitions, and attitudes have been presented in the teamwork literature, this article presents and briefly describes only a subset of those "components of teamwork" previously identified as critically important.

Team members may exhibit several behaviors that comprise teamwork: mutual performance monitoring, backup behavior, adaptability, closed-loop communication, and leadership. Mutual performance monitoring involves monitoring fellow team members' performance while still maintaining individual responsibilities. The intent is to improve overall group performance by detecting errors, deficiencies, and overloads that inhibit the team's actions. As team members monitor their teammates' performance and are able to detect deficiencies or overloads, they are able to shift work responsibilities to other team members or provide feedback about errors if and when it becomes necessary. This is referred to as

backup behavior. Taken together, mutual performance monitoring and backup behavior result in a team that is adaptable and flexible. Adaptability has been described as the capacity to recognize deviations from expected action and to readjust those actions accordingly.

Closed-loop communication also facilitates teamwork by ensuring that accurate information is given and understood by team members. It involves the sender initiating a message, the receiver accepting and acknowledging the message, and the sender ensuring that the intended message was received. When communication fails, implicit coordination and the shared understanding of the team's task may deteriorate. Finally, effective team leaders facilitate teamwork by consistently encouraging mutual performance monitoring, backup behavior, adaptability, and communication. The team leader's failure to guide the team in these behaviors may inhibit coordinated and adaptive action and can be a key factor in deficient team performance. However, the behaviors described previously are only a part of the overall picture of teamwork. There are also cognitive and attitudinal components of teamwork that facilitate these behaviors, and a few of them are briefly described in what follows.

Shared mental models have been a frequently discussed cognitive component of teamwork. In 1995, Cannon-Bowers and colleagues suggested that shared mental models are shared understandings or representations of the goals of the team, individual team member tasks, and how the team members will coordinate to achieve their common goals. Similarly, familiarity of task-specific teammate characteristics is also an important component of teamwork. Familiarity includes knowledge of the task-related competencies, preferences, tendencies, and strengths and weaknesses of teammates. Knowledge of teammates and their weaknesses can have a profound influence on team communication and interaction patterns. In fact, research has demonstrated a moderate relationship between high levels of familiarity in teams and both productivity and decision-making effectiveness.

The primary attitudinal component of teamwork is team orientation. Team orientation is frequently described as a preference for working with others. More specifically, however, it is a tendency to enhance individual performance by placing value on group input and using input from other group members while performing a group task. When discrepancies occur among team members' levels of understanding, those with a team orientation have been found to spend more time comparing their comprehension of the situation with that of their teammates to resolve the discrepancy. This provides greater opportunity to reevaluate their shared understanding of a task and to assist in error detection. Team members who do not have a team orientation are more likely to reject the validity of input from other team members, and this can potentially lead to a decline in team performance. Not all employees see value in working with others, and those who do not are less likely to succeed on teams or to be as cooperative as others. Research has shown that better decisions and higher performance result from teams that contain team-oriented individuals.

Now that the primary cognitive, behavioral, and attitudinal components of teamwork have been examined, the next section presents some concrete explanations for why teams might go off-track or "derail." Although other factors can lead to deficient teamwork (e.g., individual differences, organizational climate, stress), the section focuses specifically on factors that can be influenced through team training and team building.

3. TEAMWORK DERAILERS

Teamwork does not "just happen." Often, an organization will place a group of well-qualified individuals together to work on a project and is surprised when the team does not achieve greatness. Despite the many benefits of teamwork and cooperation at work, teamwork does not occur automatically. But why do teams derail, and how does teamwork erode? Team derailment is defined as the process whereby an effective team experiences a significant decline in performance. This section, from the vantage point of a team competency framework, identifies three primary culprits: high levels of stress; lack of knowledge of the team's mission, objectives, norms, and resources; and dissolution of trust. Any of these may lead to a decrement in team performance and cooperation.

Stress can have a significant effect on an individual's or a team's ability to perform a job. For example, teams operating in the military often face a high-stress environment in which multiple tasks must be performed under time pressure and in ambiguous situations. However, stress can result from situations in which resources are stretched or exceeded, there is danger of physical or personal harm, or the capacity to perform is diminished. Outcomes of stress for individuals in teams can be physiological (e.g., increased heartbeat, sweating), emotional (e.g., fear or anxiety), social (e.g., difficulty in coordinating actions), cognitive (e.g., a narrowing of attention), or have a direct influence on performance (e.g., trade-offs in the speed and accuracy of task performance, impaired decision making). None of these

outcomes is particularly pleasant, and all can potentially impair team performance. One solution for teams operating in high-stress environments may be to provide stress exposure training (SET), which is designed to reduce stress through the provision of information, skills training, and practice to trainees.

To avoid team derailment, team members also need to know their team's strategic and tactical goals, what resources are available to help achieve those goals, and under what conditions the team members will be expected to perform. When this knowledge is lacking, coordination of team member actions and provision of backup behavior and support are also deficient. It has been suggested that when a team initially forms, team members should be given the opportunity to gauge each other's abilities and skills and to cooperatively determine the purpose of their team. This is often referred to as creating a shared mental model of the team, the task, and the processes by which the team will complete the task. Having completed this, the team can then determine how it will accomplish its goals and establish norms of behavior and performance. When a team lacks a clear mental model, it fails to recognize and integrate task contingencies as they arise. Similarly, failure to maintain a shared perception of the situation, strategic and tactical goals, and the resources available to accomplish those goals can lead to decreased cooperation and teamwork.

Finally, trust has become a common construct in the teamwork literature due to its positive relationship to organizational citizenship behaviors, successful negotiation, and conflict management. Understandably, all of these behaviors are essential to cooperation as well. In 1998, Rousseau and colleagues defined trust as a willingness to accept vulnerability due to positive expectations of the intentions or behaviors of others. Mutual trust in a team is a deeply held confidence toward the team's climate, internal environment, and members. It is an attitude held by team members regarding the climate or mood of the team's internal environment. When team members work interdependently to achieve group goals, they must be willing to accept a certain amount of risk so as to rely on each other and openly share information. Lack of trust in this situation may result in a reduction of information sharing as well as extensive time spent in monitoring team members, not for the purpose of improving performance but rather for ensuring their trustworthiness. When organizations fail to create policies that foster trust among team members or do not allow sufficient team interaction, cooperation will deteriorate. Now that some ways in which teams and teamwork may erode have been presented, the next section describes ways in which teamwork may be enhanced in work organizations.

4. FACILITATORS OF TEAMWORK

A successful team is infused with an energizing spirit that draws the participants together into a cohesive unit and has everyone pulling together to reach a common goal. In general, teamwork can be facilitated, promoted, or encouraged at either the organizational or team level. Important organizational variables include the amount of organizational support offered and the design of the team (e.g., size, diversity, skill composition). At the team level, effective leadership, closed-loop communication, interdependence, and cohesion all contribute to team effectiveness. Although this is clearly not an exhaustive list of all elements that contribute to team functioning, these variables are instrumental in facilitating effective team performance. Each is briefly discussed in turn.

Organizational support is a primary facilitator of teamwork. It is the organization and managers that must instill cultures that encourage common objectives, shared values, mutual trust, frequent and honest communication, and the ability to act in a unitary manner. Sometimes, organizations maintain cultures that do not promote interaction. Instead, these organizations maintain a culture in which blame is constantly placed and responsibility is shifted elsewhere. When this happens, teamwork cannot be facilitated. Furthermore, it is the responsibility of the team leaders to structure team tasks. This includes delegating and assigning the tasks, managing resources, and establishing priorities. Broadly speaking, team leaders set the direction and provide the resources, whereas team members enact organizational objectives and engage in teamwork in service of their organizations.

The design of the team is another organizational level factor that can act as a facilitator of teamwork. Factors such as the size of the team, the diversity within it, and team member skill composition are important. As such, proper selection of team members is important to facilitating team performance. Research suggests that effective teams are composed of between 6 and 10 members. Larger teams provide individual members with the opportunity to "social loaf" as individual inputs may become uninterpretable and indistinguishable from the group's product or service.

Team diversity is another design issue. Research has demonstrated that the amount of cooperation that occurs in teams can be dependent on the degree of heterogeneity of individuals within it, and this heterogeneity can be

related to any differences such as those in skill, qualifications, gender, and/or ethnicity. For example, accomplishing innovation and creativity goals in product teams is more likely when members are very diverse (i.e., heterogeneous). Conversely, teams whose goals include quick production and decision making exhibit increased levels of teamwork when members are very similar (i.e., homogenous). Typically, however, teams composed of heterogeneous members will take longer to begin working together effectively. Over time, and as team members become more familiar with each other, team performance generally increases for teams composed of diverse members.

Team members must display critical knowledge, skill, and attitude competencies while performing in dynamic environments. Team researchers have suggested that general team skills should be used when selecting team members to ensure team effectiveness. Ideally, individuals possessing both task-specific skills (i.e., those relating directly to the tasks to be performed) and team-contingent skills (e.g., those with a team or collective orientation) would be selected to participate in teams. Unfortunately, this luxury is not always available. So, how do the team-level variables of leadership, closed-loop communication, and shared awareness influence team performance and effectiveness?

Team leadership may be the most important element in creating a cooperative work environment. Leadership affects team effectiveness through many processes (e.g., cognitive, motivational, affective). Team leaders set the tone by articulating a clear and motivating vision and creating a supportive climate that promotes effective team processes and behaviors, which may include specific advanced planning, communication of critical process information, and promotion of functional coordination among team members. Information from all team members must be pooled to develop plans and evaluate the consequences of team decisions. However, team leaders must remain flexible and must adapt plans and team strategies to meet task demands so as to avoid a deterioration of teamwork in dynamic and stressful situations. Team leaders' failure to guide and structure team experiences that would otherwise facilitate the development and maintenance of cooperative interactions can be a key factor in ineffective performance. An important point not frequently discussed in teamwork literature is that team members cooperate and commit to teams for different reasons. Some may be motivated to cooperate due to a desire for results, others may be motivated by the satisfaction of working with others, and still others may have less admirable aims.

Therefore, leaders should not assume that motivations are the same for all team members. However, if handled appropriately, even individualistic team member motivations can contribute to common overall team goals.

Communication is also essential for teamwork. Communication creates a shared understanding of the task at hand and ensures that all members possess the required and precise information needed. This is especially important for teams operating in very complex and dynamic environments (e.g., flight deck crews, surgical teams) because it facilitates teamwork behaviors (e.g., backup behaviors) and shared mental models that allow teams to adapt quickly and implicitly. As mentioned previously, closed-loop communication is a particularly important enabler of effective teamwork. Communiqués presented in the prescribed form, tempo, and cadence are examples of proceduralized closed-loop communication used in the military today.

Finally, teamwork behaviors are facilitated by a sense of cohesion and interdependence among team members. Teamwork involves group members' collectively viewing themselves as a group whose success depends on their interaction. Cohesion can be understood as a shared value, whereby the group interests, responsibilities, and success all are taken into account, even at the expense of individual goals. Similarly, interdependence is also typically seen as an essential characteristic of team performance. Effective teams not only have an awareness of the importance of their role in team functioning but also put this into action by fostering within-team interdependence. Effective teams contain individuals who know one another's jobs, offer to help fellow members when it is appropriate to help, and depend on one another for this help.

Discussions of both the possible derailers and potential facilitators of teamwork have now been provided. However, very little has been said about how to promote these behaviors, cognitions, and attitudes. The next section describes ways in which cooperation can be promoted at work through training.

5. PROMOTING COOPERATION AND TEAMWORK AT WORK THROUGH TRAINING

The primary purpose of team training is to provide skills that enhance teamwork. For instance, team training often emphasizes interpersonal and feedback skills in addition to clarifying the need for mutual performance

monitoring, backup behavior, and goal setting. Without this training, efficient and effective taskwork might be handicapped or hindered altogether. Team training is not a program, a place, a simulator, or a collection of individuals being trained together; rather, it is a set of instructional strategies designed to enhance teamwork competencies. Nor is team training the same as team building. Team building primarily targets role clarification and has a greater influence on team member attitudes than on performance.

Although many types of instructional strategies have been developed to assist individuals in teams to cooperate at work, only four of the most prominent and relevant ones are briefly discussed in this section: cross-training, team coordination training, team self-correction training, and team leader training. A discussion of team building is also provided, with a specific focus on the distinction between team building and team training.

The first strategy, cross-training, generally involves training each individual member on the tasks of all other team members. The goal is to create shared task models and knowledge of task-specific role responsibilities to facilitate enhanced performance monitoring and backup behavior. Knowledge of the team's mission and objectives is also targeted in cross-training. This strategy is primarily information based but may also involve simulation exercises.

Team coordination training is similar to cross-training in that it creates a general understanding of teamwork skills, but it does so with a more specific focus on promoting mutual performance monitoring and backup behavior. Coordination training uses practice and behavioral modeling to teach members about teamwork skills and help them understand the underlying processes of teamwork. Team members may be given opportunities to practice and demonstrate the skills taught in low- to moderate-fidelity simulations. These are similar to the ideas that are presented in discussions of crew resource management (CRM) and are frequently used in aviation and military settings.

Another type of training that improves mutual performance monitoring, as well as initiative, leadership, and communication, is team self-guided correction training. The team directly drives this instructional strategy in that team members are taught techniques for monitoring their own behavior and assessing the degree of training effectiveness. The team members then use feedback regarding their performance and effectiveness to correct problems that may have surfaced. The primary focus of this strategy is improved communication and increased levels of performance

feedback to facilitate future performance. Feedback enables team members to perform their subtasks competently and demonstrates the contribution of that performance to the performance of others and of the team as a whole.

Finally, team leader training aims to create leaders who are better able to motivate others, structure tasks, and facilitate shared task models of teamwork behaviors within the team. The emphasis is often on teaching team leaders to effectively debrief their members using input from their whole teams regarding specific behaviors rather than general results or proxies of effectiveness. Lecture, behavioral modeling, and practice are the typical mechanisms used to deliver this type of training.

Another technique that is often used to improve team functioning and effectiveness is team building. Team building should not, however, be confused with team training. Although team training and team building are seemingly similar, they approach teamwork from different perspectives. For instance, team training is generally intended to facilitate a shared understanding of job-related KSAs, whereas team building focuses on the process of teamwork and is intended to assist individuals and groups in examining their own behavior and relationships. It is important to note, however, that team building has been found to have a greater impact on the attitudes of individual team members than on overall team performance. Researchers have suggested that this is due to the fact that team building typically focuses on interpersonal relations, communication, problem solving, goal setting, and/or decision making rather than focusing directly on affecting performance. Because team building is generally one-dimensional in its aim (e.g., problem solving, interpersonal skills), its impact on overall team performance is limited. Finally, it should also be noted that team building is not a one-shot deal; observable improvements in cooperation may take time to appear and often require ongoing attention to attain long-term effectiveness.

Taken together, the importance of team training (and, to a lesser extent, team building) is that it focuses on deficiencies in teamwork KSAs to bolster and coordinate team member actions. Both team and organizational performance are likely to improve as a result of team training interventions. Studies conducted in military and aviation settings have found that enhanced training has resulted in significant team performance improvements. For example, researchers have seen as much as a 45% improvement in mission performance following team training. This includes a 33% improvement in tactical decision-making performance and a 10 to 34%

improvement in team coordination. Regardless of the training strategy used, it is important to measure the impact of training on team performance to determine its utility to the organization. The next section discusses issues related to measuring team effectiveness.

6. MEASURING TEAM PERFORMANCE AND EFFECTIVENESS

The underlying assumption of team training is that teams can be led to perform more effectively. However, to determine whether the enhanced instruction produced its desired effect, the effectiveness of training must be assessed. The effectiveness of training refers to the degree to which desired training outcomes are achieved, that is, the degree to which a team accomplishes its assigned tasks and mission objectives. The most important of these outcomes is the degree to which trainees transfer learned material to the job. When training transfers, the organization can be certain that learned material is used by trainees when they return to their normal work roles. Team performance measures can focus on any section of the team effectiveness input, process, or outcome system and can indicate how the team carried out assigned tasks and objectives.

Extant literature on teams and team effectiveness has identified a number of issues that should be addressed during measurement. Specifically, Salas and colleagues identified five ingredients for any team performance measurement system. First, it must be theoretically based. Second, it must consider multiple levels of measurement. Third, it must be able to describe, evaluate, and diagnose performance. Fourth, it must provide a basis for remediation. Fifth, it must support ease of use. These requirements are described in turn.

The measurement of team performance must be theoretically based, preferably on previously identified models of teamwork. After all, theory building requires that established models be empirically tested to advance the current state of knowledge. Capturing, defining, and measuring the interactions that characterize teamwork are basic to our ability to understand its meaning. Furthermore, understanding the KSAs that define teamwork is critical to establishing a nomological net of the interrelationships of these variables that should be used as the basis for structuring measurement tools. Without grounding the measurement of team performance in theoretical models of

teamwork, team theory cannot move beyond the conceptual stage.

The level of measurement must also be considered. Multiple levels of measurement should be used in the measurement system because teamwork occurs at multiple levels (e.g., individual vs team, process vs outcome). To begin, team performance needs to be evaluated at the individual level, as well as at the team level, to get a complete picture. Effective team performance can be enabled by the actions of team members operating both individually and as a collective.

Often it is impossible to determine what part of performance is due to an individual's abilities and what part is due to the actions of the team as a whole. For teams with a large number of members, individual input to the overall task may be obscured, causing individuals to lose sight of team goals. When this happens, team members will often put their needs in front of helping other team members to achieve the collective goal. Similarly, research has suggested that many individuals prefer individual assessment and find that method to be fairer. This is especially the case when individuals perceive that their individual performance is higher than what is observed at the team level. The consensus seems to indicate that individual performance should be evaluated throughout, whereas team-level outcomes should be assessed following completion of the training task or objective.

Another issue concerning the level of measurement relates to whether a greater emphasis should be placed on evaluating teamwork processes (e.g., compensatory behavior, communication, information exchange, shared task models) or team outcomes (e.g., accuracy, timeliness, product quality, goal accomplishment). Although the outcome of performance is nearly always significant, the measurement instrument should clearly differentiate between the outcome and how the outcome was achieved. In team performance measurement, the decision to evaluate team effectiveness as a team process or a team outcome is particularly important when the purpose of performance evaluation is to provide feedback for performance improvement.

Although many managers focus on team outcomes (e.g., factors that directly affect the bottom line), this evaluation provides no information to the team as to how to make performance improvements. For instance, telling sales team members they did not make the goal of selling 300 widgets provides less information to guide performance improvement than does telling the same team that there had been a noticeable decrease in information exchange among the team members during the past month that resulted in

decreased sales. Measurement must provide information that indicates why processes occurred and how those processes are linked to performance outcomes.

To be of any value, the measurement mechanism must be able to describe, evaluate, and diagnose performance. Measures must be sensitive enough to document the moment-to-moment interactions and changes in performance. Measures must also distinguish between effective and ineffective processes, strategies, and teamwork behaviors. The level of evaluation, type of performance to be evaluated (i.e., process or outcome), and environmental constraints within the organization collectively determine the measurement tools needed to measure team effectiveness. However, if the performance evaluation is conducted for developmental or remediation purposes, it is especially important that the tool be able to describe, evaluate, diagnose, and provide feedback regarding performance. Cannon-Bowers and colleagues provided a concise framework of team performance measurement tools. A few of the more flexible tools (i.e., those that can be used in most situations) are observational scales and expert ratings. More situation-specific tools include decision analysis, policy capturing, and archival records. In a broad sense, performance is typically evaluated through direct observation (e.g., employee behavior) or indirect observation (e.g., employee records, customer communication). The decision to use direct or indirect observation generally depends on what information is available to the evaluator. Both forms have their pros and cons. Similarly, the choice of measurement tool may be guided by the particular organizational and situational constraints imposed on the researcher.

Next, team performance measurement systems must outline a path to improvement. From the performance diagnosis generated, a logical basis for remediation that includes feedback, knowledge of results, and direction for subsequent instruction must be provided. Practitioners must be able to provide teams with feedback necessary to improve future performance. Measurement tools must provide information that indicates why processes occurred as they did and how particular processes are linked to certain outcomes.

A final ingredient of team performance measurement systems is that they must be easy to use. The focus of a particular measure should always be operationalized from a construct that is relatively free of contamination and deficiency, and it should not be tedious or cost-prohibitive to use. One challenge of team research has been to develop powerful, reliable, and valid measurement techniques for measuring team performance.

A final topic to consider within team measurement should be noted. That is the idea of team mental models or the "shared cognition" of a team. The idea that shared mental models allow team members to coordinate by anticipating and predicting each other's needs through common expectations of performance was discussed previously. However, the ability to measure these expectations as an external observer has proven to be more difficult than have other team processes. Currently, indirect measures of shared cognitions measure individual mental models of the team and task and aggregate the individuals' mental models to assess the degree of overlap between the cognitions used. Further research is needed to develop tools that can directly address team shared cognition. The diversification of techniques such as automated measures, or measures that are embedded in the task (e.g., simulations), also provides avenues for future research endeavors.

7. CONCLUSION

Experience has shown that teamwork and cooperation can substantially affect organizational effectiveness. However, organizations, managers, executives, and leaders need to know what develops and fosters teamwork as well as what derails it. This article has discussed why teamwork is important.

The article also discussed the nature of teamwork, indicating that it consists of the KSAs that are displayed in support of one's teammates, objectives, and mission. Ways in which cooperation and teamwork can be derailed were examined as well. Similarly, a presentation of teamwork facilitators was provided. For example, very often it is the responsibility of the leader to create a shared understanding of the vision and goals of those who report to him or her. Furthermore, it is the leader who is usually able to obtain resources and distribute them appropriately, whether through reward systems, work tools, personnel, or training opportunities. At the group level, many factors contribute to the level of cooperation that occurs among individuals. Not only do shared vision, the task itself, and the organizational culture play a part, but factors such as the level of diversity, familiarity, and the number of individuals working together also affect cooperation.

The article also presented a discussion of how training is used to promote cooperation and teamwork. Training opportunities are especially important from a cooperation perspective in that training teamwork skills is imperative to effective cooperation. Training

allows members to gain knowledge about their teammates' skills, abilities, roles, and responsibilities, and this allows them to better anticipate each other's needs and to adapt to the changing organizational environment and tasks. Training can also assist in the creation of a shared vision through goal setting and interpersonal relations.

The article then examined the measurement requirements for team effectiveness research. The fact that teamwork and training effectiveness must be measured should not be lost on the reader. Without effective description, evaluation, and diagnosis of team effectiveness, we would not know where to begin to correct performance deficiencies.

From an organizational perspective, teams and the social system in which they are embedded are interdependent, such that a change in one will affect the other. Thus, organizations that encourage cooperative behaviors through their leadership and organizational systems (i.e., reward systems) will facilitate and encourage cooperation among employees. There is clear evidence that cooperative work arrangements improve overall organizational performance. Individuals also have a role in the effectiveness of cooperation and teamwork as an organizational tool. Evidence indicates that attitudinal factors such as team identity, orientation, commitment, and trust are important. Commitment to the group will increase the willingness of the individual to expend effort and have higher involvement in the group task that is required for cooperation.

Understanding the nature of teamwork and how its interrelated components manifest themselves in terms of cognitions, behaviors, and feelings is critical to promoting coordinated adaptive team performance in increasingly complex and dynamic environments.

Thus, actions taken to improve cooperation and teamwork can have a substantial impact on the organizational bottom line.

See Also the Following Articles

Organizational Structure ■ Work Teams

Further Reading

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo, & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). San Francisco: Jossey–Bass.

Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. *American Psychologist*, *54*, 129–139.

Jones, G. R., & George, J. M. (1998). The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of Management Review*, *23*, 531–546.

McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. Guzzo, & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 149–203). San Francisco: Jossey–Bass.

Rothschild, J., & Whitt, J. A. (1986). *The cooperative workplace: Potentials and dilemmas of organizational democracy and participation.* New York: Cambridge University Press.

Salas, E., & Cannon-Bowers, J. A. (2000). The anatomy of team training. In S. Tobias, & J. D. Fletcher (Eds.), *Training and retraining: A handbook for businesses, industry, government, and military* (pp. 312–335). New York: Macmillan.

Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups: From the Hawthorne studies to work teams of the 1990's and beyond. *Group Dynamics*, *4*, 44–67.

Coping

Carolyn M. Aldwin and Loriena A. Yancura

University of California, Davis, California, USA

GLOSSARY

adaptive defense mechanisms Coping behaviors, including humor and altruism, that help to alleviate anxiety; also called mature defense mechanisms.

coping strategies Consciously chosen, intentional behaviors and cognitions that can flexibly respond to environmental demands.

coping styles Consistent methods of dealing with stress that focus on information accessing as well as emotions.

defense mechanisms Unconscious dispositional behaviors (e.g., defensive styles) associated with pathology.

depersonalization The act of separating one's identity or sense of reality from a situation.

emotion-focused coping A wide range of strategies (e.g. avoidance, withdrawal, expressing emotion, and the use of substances such as alcohol or food) that are directed toward managing one's emotional response to a problem.

meaning making A coping strategy that involves trying to see the positive or meaningful aspects of the stressful situation, especially with severe or chronic stressors. It is sometimes referred to as cognitive reappraisal.

problem-focused coping Behaviors and cognitions aimed at solving a problem that is generating anxiety. It is also called instrumental action.

psychotic defense mechanisms Individual distortions of a situation that may involve outright denial of reality, neurotic defense mechanisms, or may consist of denial of the significance of an event, such as intellectualization. They are sometimes called immature defense mechanisms.

religious coping A multifaceted coping strategy including elements of social support coping, problem-focused coping, and emotion-focused coping, that aims to conserve or transform meaning in the face of adversity.

sequelae Pathological condition and/or a secondary consequence, in this case resulting from trauma.

social support coping Seeking advice, emotional aid, and/or concrete aid from others.

In the past 25 years, more than 30,000 articles have been published on coping. This is not only a testament to the popularity of this construct but also reflects the enormous interest in the psychological and physical effects of stress. Theoretically, at least part of the reason for differential vulnerability and resilience to stress lies in how individuals cope with stress. Testing this hypothesis has led to the growing understanding of the complexity of the adaptation process. There is simply no "silver bullet" that will magically solve all problems but, rather, the effects of coping vary by how coping strategies are assessed, the type of problem, and a variety of situational factors, as well as the skill and resources of the individual facing the problem.

1. TYPES OF COPING

There are three major approaches to the study of coping: psychoanalytic approaches, coping styles, and coping

processes. These approaches to coping differ as to whether they focus on unconscious mechanisms, personality or information processing styles, or environmental influences.

1.1. Psychoanalytic Approaches

Psychoanalytic approaches focus on the use of defense mechanisms, which are unconscious ways of fending off anxiety. Defense mechanisms differ from coping strategies in that they are unconscious, dispositional (e.g., defensive styles), and are associated with pathology, whereas coping strategies are consciously chosen, intentional, and can flexibly respond to environmental demands. However, defense mechanisms vary in the degree to which they distort reality and may be hierarchically arranged from psychotic to adaptive or "mature." Psychotic defense mechanisms may involve outright denial of reality, whereas "immature" or neurotic mechanisms may consist of denial of the significance of an event, such as intellectualization. The most adaptive or mature defense mechanisms include humor and altruism.

Because they are believed to act on an unconscious level, defense mechanisms are difficult to study and they are usually assessed using in-depth interviews and case studies. Thus, there have been few large-scale studies of their effectiveness. However, they did provide the initial impetus for the study of how individuals cope with stress.

1.2. Coping Styles

Coping styles differ from defense mechanisms in that they focus on how people deal with information as well as emotions, and are thought to be consistent across time and situation. As such, they typically take the form of dichotomies such as repression–sensitization, blunting–monitoring, and approach–avoidance. Research studies have typically found that approach styles are associated with better outcomes than avoidance styles. However, this does not hold true across all circumstances, such as with uncontrollable stressors. Furthermore, studies that empirically test consistency in coping across time or situations are rare, and the ones that do include such tests generally find that individuals are not very consistent in how they cope with different types of problems.

1.3. Coping Processes

In contrast to the two previous views, the coping process approach argues that coping is responsive to both environmental demands and personal factors such as values and beliefs. As such, coping strategies are flexible and unfold over time, either in response to changing appraisals or as a function of developmental processes. Coping strategies are proactive and are not simply responses to environmental contingencies. One can appraise a situation as benign, in which case no coping is needed, or as involving threat, harm, or loss or as a challenge, all of which may evoke various coping strategies. Because this approach emphasizes the flexible nature of coping, the focus is on how individuals cope with particular situations and, as mentioned earlier, several studies have shown that coping strategies do vary across situations.

There are many different conceptualizations of coping strategies, but the five general types of coping strategies are problem-focused coping, emotion-focused coping, social support, religious coping, and meaning making. Problem-focused coping, also called instrumental action, encompasses behaviors and cognitions aimed at solving the problem, such as seeking information, taking direct action, or breaking the problem down into more manageable pieces, a strategy referred to as "chunking." Sometimes, delaying or suppressing action can be a useful problem-focused strategy. For example, purposefully delaying a direct confrontation with someone may lead to a more rapid solution to a problem than acting in anger.

Emotion-focused coping includes a wide range of strategies that are directed toward managing one's emotional response to the problem. Some examples are avoidance, withdrawal, expressing emotion, and the use of substances such as alcohol or food. As might be expected, avoidance strategies are often associated with poor outcomes, but other emotion-focused strategies, such as expressing emotion through journals or writing, may be associated with positive outcomes.

Social support involves seeking both emotional and concrete aid from others or advice. The outcome of these types of coping strategies often depends on the social context. For example, confiding in others after a trauma is generally associated with better outcomes, but if the confidant responds negatively, emotional distress may be increased.

The study of religious coping is relatively new. It can contain elements of social support or problem-focused and emotion-focused coping, and it seeks to conserve or transform meaning in the face of adversity. In general, religious coping is associated with positive outcomes, but it does have negative guises. Belief in a punitive God, or feeling that one has been treated unfairly or been deserted, may be associated with much poorer outcomes.

The final category, meaning making, is the least well understood. It is sometimes referred to as "cognitive reappraisal" and involves trying to see the positive or meaningful aspects of the situation, especially with severe or chronic stressors. As with religious coping, how one goes about making meaning may affect its association with outcomes. Simply asking "Why me?" may be associated with poorer outcomes, but coming to realize how a problem fits into the larger pattern of one's life may be a painful process but in the end may be one way in which individuals grow from stressful or traumatic experiences.

1.4. Measurement Issues

Coping style measures ask individuals how they typically cope with problems. The dichotomized nature of coping styles lends itself well to psychometric assessment. The factor structures of coping style inventories are generally stable, and the factors correlate well with psychological outcomes. Given that coping may not be very stable, however, their ecological validity is somewhat questionable.

Process measures ask individuals to select a recent problem and indicate how they coped with that type of problem. Unfortunately, process measures typically have somewhat "messy" psychometrics, and their factor structures may be unstable. This makes comparison across situations or over time problematic.

Sometimes a more accurate picture can be obtained using multiple assessments, either through daily diaries or through event sampling throughout the day. Although some scholars have criticized the psychometric properties of these measures (often consisting of only one item per strategy), these types of procedures may be especially useful in the context of coping with chronic stressors because they permit within-subject analyses that may yield strong relationships between coping and outcomes.

2. COPING AND HEALTH OUTCOMES

2.1. Mental Health Outcomes

In many ways, being able to successfully cope with problems can be considered a hallmark of psychological health. Difficulties in coping may be associated with emotional distress, including depression and anxiety. However, the relationship is neither simple nor straightforward. Because there may be differences between how people cope with trauma and with more everyday stressors, these two different types of problems are addressed separately.

2.1.1. Coping with Stressors

The relationship between coping and psychological symptoms is characterized by complexity along a number of different dimensions. First, coping strategies have situation-specific effects. Many studies have demonstrated that the effectiveness of coping strategies rests on the controllability of the stressor. Controllable stressors, such as problems at work, are more effectively managed with problem-focused coping, whereas uncontrollable stressors, such as the death of a loved one, are more effectively managed with emotion-focused coping.

Second, there may also be a confound between stress, distress, and coping. Higher levels of stress may evoke more coping strategies and also create more distress; thus, all coping strategies will correlate with higher levels of distress in particularly difficult circumstances. Only by controlling for the stressfulness of the problem can the true relationship between coping and mental health be revealed.

Third, there is a difference between short- and long-term outcomes. Avoidance coping strategies, for example, may decrease psychological distress in the short term but cause great distress in the long term if the underlying problem remains unresolved. Brief periods of respite, however, can allow individuals to "recharge their batteries" and promote better problem-focused coping in the long term.

Fourth, the causal directionality of coping may not be clear. Rather than coping resulting in depressive symptoms, individuals who are depressed may cope in different ways than nondepressed people.

Fifth, people may employ the same coping strategies with different levels of skill. Thus, it makes sense to determine the efficacy of the strategy—that is, whether it "worked" for that person in that particular situation.

Finally, most research has focused only on coping as a means of reducing symptoms. However, effective coping may lead to more positive outcomes, such as increases in mastery, self-knowledge, and better social relationships. Thus, it is not surprising that the literature on coping and mental health is rich but somewhat contradictory.

2.1.2. Coping with Trauma

Traumatic stressors are ones that either threaten or expose individuals to loss of life and major physical damage to

themselves, loved ones, or large numbers of individuals. Surprisingly, the trauma literature is much clearer in demonstrating the efficacy of coping strategies. In general, how individuals cope with trauma and its sequelae are better predictors of the development of posttraumatic stress disorder (PTSD) than simply exposure to trauma.

There are four ways in which coping with trauma differs from coping with less severe stressors. First, traumatic circumstances are often uncontrollable and many people fall back on the use of unconscious defense mechanisms, such as denial and repression. For example, depersonalization may be the only possible response to being tortured. Second, disclosure and seeking support from others appear to be especially important in traumatic situations. However, if the disclosure evokes negative responses from others, PTSD symptoms may become worse. Some individuals are relatively successful at "partitioning off" their trauma experience and rebuilding their lives.

Third, the process of coping with traumatic events and their sequelae is much longer and may take years. Finally, the coping strategy of meaning making is particularly important in dealing with trauma. Whether an individual is able to make some sense out of traumatic events can determine whether those events lead to positive or negative mental health outcomes.

2.2. Physical Health Outcomes

The evidence for the effect of coping on physical health outcomes such as pain, blood pressure, catecholamines, lipids, and immune system function is much weaker, in part because studies are relatively rare and may focus on special populations, such as heart disease patients or AIDS victims.

We identified four models by which coping may affect health outcomes. The direct effects model examines a simple correlation between coping and health outcomes, regardless of the situation or type of stressor. For example, instrumental action may be associated with better lipid or immune system profiles whether the stressor is major or minor. Second, coping may moderate or buffer the effects of stress on health outcomes. That is, there are no direct effects, but individuals under high levels of stress who use instrumental action may have lower blood pressure levels. The mediated effects models suggest that the effects of coping depend on a variable or variables intervening between coping and physical health outcomes. For example, many of the physiological effects

of coping on health might be explained through its effect on psychological distress. If an individual's coping strategies decrease negative affect, then it may be associated with better immune system profiles. The fourth model is the contextual effects model, which states that the effect of coping on health might depend on environmental context. For example, Zautra and colleagues showed that seeking social support may be effective in reducing pain for diseases such as osteoarthritis but less effective for rheumatoid arthritis, in part because the symptoms in the latter illness are less clear and thus may be more difficult for others to understand.

There is evidence for each of these models; however, none of them has received overwhelming empirical support. Clearly, much research is needed to illuminate the pathways by which coping influences physical health outcomes.

See Also the Following Articles

Health and Culture ■ Posttraumatic Disorders ■ Stress

Further Reading

Aldwin, C. M. (1999). *Stress, coping, and development: An integrative approach*. New York: Guilford.

Cramer, P. (2000). Defense mechanisms in psychology today: Further processes for adaptation. *American Psychologist*, 55, 637–646.

de Ridder, D., & Kerssens, J. (2003). Owing to the force of circumstances? The impact of situational features and personal characteristics on coping patterns across situations. *Psychology & Health*, 18(2), 217–236.

Folkman, S., & Moskowitz (1994). Coping: Pitfalls and promise. *Annual Review of Psychology*, 55, 745–774.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal & coping*. New York: Springer-Verlag.

Mikulincer, M., & Florian, V. (1996). Coping and adaptation to trauma and loss. In M. Zeidner, & N. S. Endler (Eds.), *Handbook of coping: Theory, research, applications* (pp. 554–572). New York: Wiley.

Pargament, K. (1997). *The psychology of religion and coping: Theory, research and practice*. New York: Guilford.

Skinner, E. A., Altman, J., & Sherwood, H. (2003). Searching for the structure of coping: A review and critique of category systems for classifying ways of coping. *Psychological Bulletin*, 129, 216–269.

Tennen, H., Affleck, G., Armeli, S., & Carney, M. A. (2000). A daily process approach to coping: Linking theory, research, and practice. *American Psychologist*, 55, 626–636.

Valliant, G. E. (1993). *The wisdom of the ego*. Cambridge, MA: Harvard University Press.

Counseling and Culture

Paul B. Pedersen
University of Hawaii, Honolulu, Hawaii, USA

global context, whether that perspective is held in the Western or Eastern geographic hemisphere.

Counseling is a product of Western cultures whose values have continued to dominate the counseling profession through the imposition of cultural assumptions as counseling services have spread to other cultures, demonstrating the generic centrality of culture to achieve competence in the counseling process.

GLOSSARY

cultural encapsulation The tendency of people to build walls or capsules around themselves by simplifying or minimizing the importance of other people's culture's.

cultural grid A framework combining personal factors (e.g., behavior, expectation, values) on one dimension and the cultural contexts in which those personal factors were learned on the other dimension.

culture-centered perspective A perspective that makes the cultural context central to an understanding of every situation.

fourth force The view of multiculturalism as a "fourth dimension" of counseling to supplement the psychodynamic, humanistic, and behavioral dimensions.

multicultural competencies A list of 34 specific competencies describing how an understanding of culture is basic to the counseling process; this list has been evolving since 1962.

Western culture A dominant culture sociopolitical perspective favoring individualism and linear thinking in the

1. INTRODUCTION

Because all behaviors are learned and displayed in a cultural context, accurate assessment, meaningful understanding, and appropriate intervention require consideration of the cultural context when providing counseling services. Making the consumer's cultural context central rather than marginal when providing psychodynamic, behavioral, and humanistic services strengthens those traditional counseling perspectives. However, because counseling first developed as a field in a Euro-American cultural context, counseling typically reflects Euro-American assumptions in textbooks, theories, tests, ethical guidelines, methods, and other services. The purpose of this article is to discuss the generic importance of making cultural context central rather than marginal to the counseling process.

2. CULTURAL BIAS IN COUNSELING

In 1996, a report by the Basic Behavioral Science Task Force of the National Advisory Mental Health Council (NAMHC) documented the extent of cultural encapsulation of mental health services. First, anthropological and cross-cultural research has demonstrated that cultural beliefs influence the diagnosis and treatment of mental illness. Second, the diagnosis of mental illness differs across cultures. Third, research has revealed differences in how individuals express symptoms in different cultural contexts. Fourth, culturally biased variations in diagnosis vary according to the diagnostic categories relevant to the majority population. Fifth, most providers come from a majority culture, whereas most clients are members of minority cultures. If the standard practices of mental health services are themselves encapsulated, as suggested by the NAMHC report, these cultural biases will certainly influence the practical applications of psychology through counseling and therapy.

Counseling has typically assumed that there is a fixed state of mind that is obscured by cultural distortions and that relate behaviors across cultures to some universal definition of normative behavior described in textbooks. A contrasting anthropological perspective assumes that cultural differences are clues to divergent attitudes, values, or perspectives that differentiate one culture from another based on a culture-specific viewpoint. Anthropologists have tended to take a relativistic position when classifying and interpreting behavior across cultures. Psychologists, in contrast, have linked social characteristics and psychological phenomena with minimal attention to cultural differences. When counselors have assumed the same interpretation of similar behaviors regardless of the cultural context, cultural bias has been the consequence. The influence of cultural bias has led to the encapsulation of counselors by dominant culture perspectives regardless of the cultural context.

3. CULTURAL ENCAPSULATION OF COUNSELORS

In 2001, Ponterotto and colleagues described how counseling has a reputation among many minorities as protecting the status quo of the dominant culture by requiring minority clients to fit the dominant culture system even when that system is acknowledged to be unjust. In 1999, Sue and Sue described how these attitudes have resulted in documented examples of scientific racism.

In 1962, Wrenn first introduced the concept of cultural encapsulation. This perspective assumes five basic identifying features. First, reality is defined according to one set of cultural assumptions. Second, people become insensitive to cultural variations among individuals and assume that their own view is the only right one. Third, assumptions are not dependent on reasonable proof or rational consistency but are believed to be true regardless of evidence to the contrary. Fourth, solutions are sought in technique-oriented strategies and quick or simple remedies. Fifth, everyone is judged from the viewpoint of one's self-reference criteria without regard for the other person's separate cultural context. The encapsulation has not diminished over time. There is evidence that the profession of counseling is even more encapsulated now than it was when Wrenn wrote his original article.

In 1994, Albee described how completely psychology—and counseling as based on a psychological foundation—in the United States has been encapsulated during the past 100 years:

> Most of the early leaders in psychology embraced ideological views that stressed the natural superiority of a white male patriarchy, the acceptance of social Darwinism, [and] the inferiority of women and of the brunette races. Calvinism stressed economic success as the hallmark of salvation, and psychology concurred. Anti-semitism and homophobia were standard. Eugenics spokesmen urged the elimination of the unfit and inferior and opposed welfare programs, decent wages, and safe working conditions.

These views continue to be held today, although sometimes in more subtle forms. The encapsulation of counselors is most evident in the way in which counselors are trained and educated.

4. CULTURALLY LEARNED ASSUMPTIONS IN COUNSELOR EDUCATION

The training of counselors has been characterized by conventional assumptions that reflect the values and priorities of the Euro-American cultural context. Table I describes 10 examples of these assumptions that Pedersen used in 2002 to demonstrate their pervasive influence in applications of counseling.

TABLE I
Western Cultural Assumptions in Counseling

Euro-American cultural assumption	*Comments*
Counselors all share the same unimodal measure of what is normal behavior.	There is a frequent assumption that describing a person's behavior as "normal" reflects a judgment that is both meaningful and representative of a desired pattern of culture-specific behaviors across social, economic, political, and cultural contexts, that is, represents the dominant culture perspective of culture-specific behavior.
Individuals are the basic building blocks of society.	This individualistic presumption is reflected in psychological terms such as personality and self-awareness. Conventional counseling is focused primarily on the development of individuals rather than collectivities or groups such as the family, the organization, or society itself.
Only problems defined within the framework of the counselor's specialized expertise or academic discipline are of proper professional concern to the counselor.	There is a tendency to isolate the professional identity of the counselor from that of other professionals even though multicultural problems wander across these boundaries freely.
There is a superior quality judgment attached to psychological abstractions.	In the counselor's use of professional jargon, he or she assumes that the same words have the same meaning across cultural contexts. Although this assumption is typical of a "low-context" culture where the context is less relevant, it would not apply to "high-context" cultures where all meaning is contextually mediated.
Independence is usually desirable, and dependence is usually not desirable.	As part of the Western emphasis on individualism, there is a belief that individuals should not be dependent on others or allow others to become dependent on them. This is not the case in a more collectivist culture.
Clients are helped more by formal and professional counseling than by their natural support systems.	Family and peer support are the primary resources in many cultures where professional counseling is less available or is inappropriate. In some non-Western cultures, counseling is a last resort only when everything else has failed.
Everyone thinks the same way, moving linearly from cause to effect.	It is not just the content of one's thinking that is culturally mediated but also the very process of thinking itself. Nonlinear thinking, typical of many non-Western cultures, will seem illogical to linear thinkers.
Counselors need to change clients to fit the system and not change the system to fit the clients.	Advocacy by counselors is frequently considered unethical and unprofessional. Much of the counseling literature relates to client adjustment, sometimes even when the system is wrong and the client is right.
History is not relevant for a proper understanding of here-and-now problems.	Counselors are more likely to focus on the immediate here-and-now events that created a crisis and to consider historical background as a distraction at best and as a defensive evasion at worst.
We already know all of our culturally learned assumptions.	Each time counselors discover one of their cultural biases, they disprove this assumption. As they increase their contact with persons and groups from other cultures, this process of self-discovery is accelerated.

One way in which to challenge the universal application of dominant culture perspectives, regardless of the cultural context, is to separate behaviors from expectations in counseling. This will demonstrate how shared positive expectations (e.g., safety, respect, fairness) might be expressed by very different behaviors and how similar behaviors (e.g., direct/indirect, loud/quiet) across cultures might be linked to very different expectations in a two-dimensional "cultural grid."

5. THE CULTURAL GRID LINKING COUNSELING AND CULTURE

The connection between counseling and culture can be described visually in a cultural grid that includes individual behavior, expectations, and values on one dimension and the "culture teachers" of those individual characteristics on the other dimension. The cultural grid, as described by Pedersen in 2000, is an attempt to demonstrate the personal–cultural orientation of the individual in relationship to the many different cultures to which that individual belongs. The cultural grid shows an "inside-the-person" or intrapersonal cultural framework for demonstrating how cultural and personal factors interact. Through the cultural grid, each individual behavior can be linked to the culturally learned expectation on which that behavior is based, and each expectation can be linked to the culture teachers from whom those expectations were learned.

The cultural grid provides a systematic framework for the counselor to interpret each client behavior in the context where that behavior was learned and is displayed. The cultural grid demonstrates that there are many different culture teachers present in the counseling interview, with some belonging to the counselor and some belonging to the client. The cultural grid further provides a framework for comprehending the client's culturally learned behavior from his or her point of view.

Pedersen pointed out that each cultural context is complicated and dynamic, influenced by many culture teachers who take turns being salient according to every time and place. An awareness of one's cultural identity requires the ability to identify how each behavior is an expression of specific culturally learned expectations and how each expectation was learned from specific culture teachers. The culture teachers might come from family relationships, power relationships, or other memberships. Judging other people's behavior out of context or without regard to why they engaged in the behavior or where they learned the behavior is likely to be misleading at best and totally wrong at worst. It is not easy to discover why people do what they do, but that does not excuse the counselor from the obligation to try. Unless each client behavior is understood in context, that behavior is very likely to be misunderstood by the counselor (Fig. 1).

The culturally learned assumptions on which current theories of counseling are based, such as those favoring individualism over collectivism, need to be reassessed so that the theories of counseling can be applied across cultures more accurately and meaningfully.

6. MULTICULTURAL THEORIES OF COUNSELING

One culture-centered perspective that applies cultural theories to the counseling process is found in a 1996 book on multicultural counseling theory (MCT) by Sue and colleagues. This approach is based on six propositions that demonstrate the fundamental importance of a culture-centered perspective:

1. Each Western or non-Western theory represents a different worldview.
2. The complex totality of interrelationships in the client–counselor experiences and the dynamic changing context must be the focus of counseling, however inconvenient that might become.
3. A counselor's or a client's racial/cultural identity will influence how problems are defined and will dictate or define appropriate counseling goals or processes.
4. The ultimate goal of a culture-centered approach is to expand the repertoire of helping responses available to counselors.
5. Conventional roles of counseling are only some of the many alternative helping roles available from a variety of cultural contexts.
6. MCT emphasizes the importance of expanding personal, family, group, and organizational consciousness in a contextual orientation.

As these MCT propositions are tested in practice, they will raise new questions about competencies of multicultural awareness, knowledge, and skill in combining cultural factors with psychological process. How does one know that a particular psychological test or theory provides valid explanations for behavior in a particular cultural context? What are the cultural

Personal Variables

	Where you learned to do it	Why you did it	What you did
Cultural Teachers			

1. Family relations
    Relatives
    Fellow countrypersons
    Ancestors
    Shared beliefs			

2. Power relationships
    Social friends
    Sponsors and mentors
    Subordinates
    Supervisors and superiors			

3. Memberships
    Coworkers
    Organizations
    Gender and age groups
    Workplace colleagues			

4. Nonfamily relationships
    Friendships
    Classmates
    Neighbors
    People like me			

FIGURE 1   Within-person cultural grid.

boundaries that prevent generalization of psychological theories and methods? Which psychological theories, tests, and methods can best be used across cultures? Which psychological theories, tests, and methods require specific cultural conditions? MCT assumes multicultural competencies.

Multicultural competencies have been developed by Sue and colleagues over the past 40 years or so. In a comprehensive description of these 34 competencies in 1998, Sue reviewed these multicultural competencies. A summary of those competencies in three dimensions, emphasizing multicultural awareness, knowledge, and skill, can be described in three developmental dimensions:

• *Dimension 1*. Counselors' awareness of their own assumptions, values, and biases includes (a) awareness of cultural heritage comfortable with differences but aware of limits; (b) knowledge about oppression, racism, and discrimination; and (c) skills in self-improvement toward a nonracist identity.

• *Dimension 2*. Understanding the worldview of the culturally different client includes (a) awareness of emotional reactions toward other racial/ethnic groups, (b) knowledge about the culture of the client population and its influence on counseling and society, and (c) skills in mental health issues of other cultures and active involvement with ethnic minority groups.

• *Dimension 3*. Developing appropriate intervention strategies and techniques includes (a) awareness of religious and spiritual indigenous mental health resources; (b) knowledge about how counseling fits with other cultures, institutions, and assessments; and (c) skills in culturally appropriate counseling of indigenous people.

These three dimensions of multicultural competency have tended to be described as abstractions, and much

more work is needed to link these abstract competencies to individual cases and situations. Ponterotto and colleagues did an excellent job of reviewing the literature already in print to apply the multicultural competencies in practice. As of this writing, Constantine and Sue are working on an even more comprehensive demonstration of how the multicultural competencies can be applied in practice. Applying these competencies in practice will reduce the number of racist incidents in the counseling profession generally.

The lack of trust in people who provide counseling services and the belief that the status quo was being held are documented in "scientific racism" and "Euro-American ethnocentrism." According to Sue and Sue's 1999 book, cultural differences were explained by some through a "genetic deficiency" model that promoted the superiority of dominant White cultures. The genetic deficiency approach was matched to a "cultural deficit" model that described minorities as deprived or disadvantaged by their culture. Minorities in the United States were underrepresented among professional counselors and therapists, the topic of culture was trivialized at professional meetings, and minority views were underrepresented populations in the research literature. Consequently, the counseling profession was discredited among minority populations because they viewed counseling as a tool to maintain the differences between those who had power and access to resources and those who did not. Making the cultural context central to counseling is in no way intended to diminish the importance of established counseling theory. In fact, making culture central as a "fourth dimension" will serve to increase the accuracy and importance of established theory.

7. MULTICULTURALISM AS A FOURTH FORCE

In 1998, Pedersen examined some of the issues involved in declaring multiculturalism as a fourth force in psychology:

1. Significant changes are taking place, even though there is disagreement about the nature of those changes.

2. Multiculturalism has become a significant domestic force in most cultures, even though it may not yet provide a coordinated global perspective.

3. Multiculturalism has sometimes been used to rationalize oppression in countries such as South Africa; consequently, it has a bad reputation in those countries.

4. It may be premature to describe multiculturalism as a fourth force, although it has had a more powerful impact in the more applied areas of psychology such as counseling.

5. The U.S. version of multiculturalism is grounded in the individualist values of that cultural context more than in a non-Western collectivist perspective.

6. Within-group differences of ethnocultural groups, such as demographics, status, and affiliation, function "like" cultures regarding issues of age, disability, gender, and other special interests.

7. Cultural similarities (e.g., among youth-age strata) probably exceed similarities across generations in each separate ethnocultural group.

8. Multiculturalism will change not only the content of our thinking but also the very process of thinking itself.

9. Making culture central enhances the meaningful usefulness of traditional psychological theories in ways that might lead us to call this new perspective a fourth dimension rather than a fourth force.

Adding a fourth dimension of multiculturalism to counseling theory is not an easy task but remains a very important aspect of making the cultural context central to counseling.

8. RESISTANCE TO MULTICULTURALISM AS A FOURTH FORCE

In 1998, Sue identified sources of resistance to the term "multiculturalism as a fourth force":

1. Some view multiculturalism as competing with already established theories of psychological explanation in ways that threaten the professions of counseling and psychology.

2. The terms "multiculturalism" and "diversity" are loosely associated with affirmative action, quotas, civil rights, discrimination, reverse discrimination, racism, sexism, political correctness, and other highly emotional terms.

3. To the extent that multiculturalism is connected with postmodernism, the arguments against postmodernism as a valid theory are also applied to multiculturalism.

4. Those favoring a universalistic perspective contend that the same practice of counseling and therapy applies equally to all populations without regard to cultural differences.

5. Others contend that there are no accepted standards for describing multiculturalism as a theory in practice and that it is too loosely defined to be taken seriously.

6. There are no measurable competencies for multicultural applications of counseling or adequate standards of practice.

7. Multiculturalism is too complicated, and it would be unrealistic to expect counselors to attend to such a range of factors simultaneously.

8. More research is needed on multicultural competencies, standards, methods, and approaches.

9. Multicultural standards cannot be incorporated into the counseling profession until all groups have been included.

10. Multiculturalism represents reverse racism, quotas, and is anti-White.

In discussing these sources of resistance, Sue pointed out the tendency to misrepresent or misunderstand the notion of multiculturalism and the dangers of that misunderstanding.

Although the task of reducing cultural bias in counseling is extremely difficult, it remains an "aspirational goal" of the profession.

9. CONCLUSION

In 1999, Sue and colleagues summarized what needs to happen for psychology programs to facilitate positive multicultural initiatives so as to include (a) faculty and student preparation in the development of cultural competence; (b) a multicultural curriculum in all aspects of education and training; (c) minority representation among students, staff, faculty, and administration; (d) an inclusive and positive campus climate; (e) recognition of culturally biased teaching and learning styles; (f) people providing a social support network and services that understand the minority experience; and (g) recognition that current programs, policies, and practices negating multicultural development must be changed.

Multiculturalism is not just a trend or fad; rather, it is a permanent phenomenon of social functioning that reflects the advanced technology of a world where international boundaries have been replaced by close interactions among people around the globe. As the primary support systems of the village and family are diminished in authority around the world, more and more societies are looking for alternatives to provide the necessary social support in their communities. The field of counseling is a promising alternative for many of the world's

cultures provided that counseling can be modified to fit with a variety of cultures and not merely impose Western or dominant culture values.

See Also the Following Articles

Cross-Cultural Psychology, Overview ■ Cultural Psychology ■ Racial and Ethnic Minorities, Counseling of ■ Training, Cross-Cultural

Further Reading

Albee, G. W. (1994). The sins of the fathers: Sexism, racism, and ethnocentrism in psychology. *International Psychologist*, *35*, 42–78.

Basic Behavioral Science Task Force of the National Advisory Mental Health Council. (1996). Basic behavioral science research for mental health: Sociocultural and environmental processes. *American Psychologist*, *51*, 722–731.

Berry, J. W., Poortinga, Y. H., Segall, H. N., & Dasen, P. J. (1992). *Cross-cultural psychology: Research and applications*. Cambridge, UK: Cambridge University Press.

Marsella, A. J. (1998). Toward a "global–community psychology": Meeting the needs of a changing world. *American Psychologist*, *53*, 1282–1291.

Pedersen, P. (1998). *Multiculturalism as a fourth force*. Philadelphia: Brunner/Mazel.

Pedersen, P. (2000a). *Handbook for developing multicultural awareness* (3rd ed.). Alexandria, VA: American Counseling Association.

Pedersen, P. (2000b). *Hidden messages in culture-centered counseling: A triad training model*. Thousand Oaks, CA: Sage.

Pedersen, P. (2002). Ethics, competence, and other professional issues in culture-centered counseling, In P. Pedersen, J. Draguns, W. Lonner, & J. Trimble (Eds.), *Counseling across cultures* (5th ed., pp. 3–28). Thousand Oaks, CA: Sage.

Ponterotto, J. G., Casas, J. M., Suzuki, L. A., & Alexander, C. M. (2001). *Handbook of multicultural counseling* (2nd ed.). Thousand Oaks, CA: Sage.

Sue, D. W., Bingham, R. P., Porche-Burke, L., & Vasquez, M. (1999). The diversification of psychology: A multicultural revolution. *American Psychologist*, *54*, 1061–1069.

Sue, D. W., Careter, R. T., Casas, J. M., Fouad, N. A., Ivey, A. E., Jensen, M., LaFromboise, T., Manese, J. E., Ponterotto, J. G., & Vazquez-Nutall, E. (1998). *Multicultural counseling competencies: Individual and organizational development*. Thousand Oaks, CA: Sage.

Sue, D. W., Ivey, A. E., & Pedersen, P. B. (1996). *Multicultural counseling theory*. Pacific Grove, CA: Brooks/Cole.

Sue, D. W., & Sue, D. (1999). *Counseling the culturally different: Theory and practice* (3rd ed.). New York: John Wiley.

Sue, S. (1998). In search of cultural competencies in psychology and counseling. *American Psychologist*, *53*, 440–448.

Sue, S. (1999). Science, ethnicity, and bias: Where have we gone wrong? *American Psychologist, 54,* 1070–1077.

Wrenn, G. (1962). The culturally encapsulated counselor. *Harvard Educational Review, 32,* 444–449.

Wrenn, G. C. (1985). Afterward: The culturally encapsulated counselor revisited. In P. Pedersen (Ed.), *Handbook of cross cultural counseling and therapy* (pp. 323–329). Westport, CT: Greenwood.

Counseling Interview

Paul H. Glasser

*Rutgers, The State University of New Jersey,
New Brunswick, New Jersey, USA*

Sara F. Fine

Bar-Ilan University, Ramat Gan, Israel

GLOSSARY

catharsis The expression of feelings and emotions by a client or patient during a clinical interview.

clinical interview A relationship between a psychiatrist, a psychologist, a social worker, or other types of professionals with a client or patient, the focus of which is the facilitation of the growth and development of the client or patient.

confidentiality A client's or patient's trust that the counselor will not reveal to others the information given to him or her.

contract An agreement between a counselor and client or patient about the goals of the treatment process and the means that will be used to achieve them.

counseling A generic term that refers to a relationship between two or more people in which one person facilitates the growth and development of others in order to help them deal with their problems more effectively.

counseling interview A dyadic relationship between a counselor and a client or patient in which the former facilitates the growth and development of the latter in order to help him or her deal more effectively with his or her problems.

intake interview The first interview by a counselor in which an assessment of the client or patient is made and an agreement is reached by both parties about what the goals of the counseling process will be and how these goals will be achieved.

presenting problem A problem the client or patient is experiencing or, when dealing with an involuntary client, the problem the referring agency defines.

privileged communication Laws that give the counselor the right to hear almost anything a client or patient says with the responsibility of not repeating it to others.

psychiatric interview A dyadic relationship between a psychiatrist and a patient for the purpose of diagnosis and treatment of a patient who has problems in interpersonal functioning and/or internal psychological and emotional difficulties and that includes the assessment and treatment of any physiological sources of the patient's distress.

trust The client or patient's belief that whatever the counselor does is in his or her best interest.

The purpose of the counseling interview is to help a client or patient with a problem he or she is experiencing or a problem the referring agency defines. This is known as the presenting problem. The counselor must be in control of the interview and respond to the client in a professional way that is different from that of friends and neighbors. The rhythm and cadence of the interview requires that the counselor allow the client to lead, focus on his or her feelings, be open and honest in response to the client, and be reflective so that the client learns that this is his or her problem. Basic to the

counseling process is that the client must learn to trust the counselor so that he or she will be open to seeing the problem differently and dealing with it more effectively. The variety of methods and techniques the counselor uses to achieve this end are described in detail in this article.

1. PURPOSE AND GOALS

The purpose of the counseling interview is to help the client with a problem he or she is experiencing or, when dealing with an involuntary client (e.g., a criminal or delinquent or a parent accused of abuse or neglect), the problem the referring agency defines. This is often called the presenting problem. Although it is imperative that the counselor deal with the problem as the client or referring agency sees it, this does not mean that the counselor always perceives the problem in the same way. For example, a wife may come to the counselor because she believes her husband does not love her anymore, but it soon becomes apparent that her interpersonal response to her husband is at least partially responsible for his negative reactions to her. However, it may take a number of sessions before the counselor believes it is possible for the client to understand the problem in a way that is different from his initial presentation. [Although both counselors and clients may be men or women, for purposes of style we refer to counselors in the feminine and to clients in the masculine.]

Early in the interviewing process, it is important to set goals for counseling and to discuss the means that will be used to achieve these ends. Agreement is reached with the client that must reflect both an understanding of the problem by the client and counselor and a way of dealing with the problem that is likely to be effective and agreeable to the client. This agreement about means and ends is called the contract. In some counseling situations, the contract is implicit in the setting. For example, the employment counselor is there to help the client find a job, or the academic counselor sees her purpose is to help the student find the best college for him. In other situations, the contract may be much more complicated; for example, the father accused of abusing his child who believes "spare the rod, spoil the child" may need to agree to consider alternative approaches to child rearing. In the latter situation, considerable discussion between the client and the counselor, sometimes over a series of interviews, may be necessary before a contract can be established.

2. DYNAMICS OF COUNSELOR-CLIENT INTERACTION

There is a rhythm and cadence in the counseling process that often needs to be established in the early interviews. First, the counselor allows the client to lead. It is the client who is asking for help and he must do so in the way that makes the most sense to him. This does not mean that the client will present material in the most logical way. Rather, he will present the "story" in his own way, which in itself should have meaning to the counselor who is listening. The organization of the material presented may have as much meaning as the content. The counselor will make note of what is not said as well as what is said. She will notice any changes in speech patterns and times of silence as the narrative flows. By allowing the client to lead, the counselor is beginning to establish the trust that the client needs to continue.

Second, the focus, particularly in the early interviews, should be on feelings. Clients come to each new counseling session, even those that seem to deal with relatively unemotional situations, with feelings—anxiety or fear, grief or despair, rage or betrayal. The student may be afraid that he will not get into the college of his choice or that he may disappoint his parents, the teenager may be furious at her parents for limiting her activities, or the husband may be overwhelmed with grief after the death of his wife. In each case, before the client can move on he must have the opportunity to express these feelings. The expression of such feelings is called catharsis. The counselor can help the client by putting an accurate name to these feelings and describing their effect, intensity, and magnitude, which helps the client to control them and then move on to deal with the problem that causes them.

A third issue is the quality of the counselor's communication. She must be open to what is heard and open and honest in her response. It means accepting the client's statements, not arguing or debating with the client about what has been said or what the client feels and not using the word "but" to rebut what the client is feeling or saying. It means being nonjudgmental even when the client expresses values incompatible with those of the counselor or that seem morally reprehensible. This does not mean that the counselor will never disagree with the client or that she approves of the client's antisocial behavior. It does mean the counselor will not do so until she has begun to establish a relationship of trust with the client, and that the

counselor will not confront the client with his unacceptable behavior until the client has sufficient positive feelings about the counselor that he can accept such confrontation with understanding and insight.

There are three therapeutic-type responses that are very useful in allowing the counseling interpersonal relationship to develop. The first we call minimal encourager. This is nothing more than nodding, murmuring "uh huh," or using simple statements such as "Tell me more" or "Please go on." Such statements are minimal and do not intrude on the client's flow of material.

The second type of response is summation of content. By summarizing or paraphrasing the client's content in her own words, the client can confirm or modify the counselor's perceptions. This allows the client to continue to control the interview situation while at the same time the counselor confirms an understanding of what is being said.

The third technique is to reflect the client's feelings about the situation or person he is describing. This is a powerful response because this teaches the client at least three things: (i) that it is okay to have feelings, no matter what they are; (ii) that the counselor is not afraid of any of his feelings, no matter how intense they are; and (iii) that naming these feelings is the first step in getting them under control. This will sometimes be followed by the client expressing a great sense of relief: "Someone finally understands." The counselor's silence after that will allow the client to think about what he said and this will often be followed by some of the most significant statements that will be made in the interview.

Finally it should be noted that the counselor's reflective responses should be such that the client learns that he owns the problem. This is particularly important if the client blames others for the problem. "My child does such terrible things I have no alternative but to physically punish him." The counselor's appropriate response is not to suggest alternative punishments at this point but rather to reflect on the client's feelings: "He must make you very angry." Thus, the mother may begin to see that it is her anger that led to her being accused of being abusive and not the child's behavior.

3. ISSUES, DILEMMAS, AND CONTRADICTIONS IN THE COUNSELING PROCESS

The counseling process is a complex one. What makes it different from other forms of interpersonal interaction is that the counselor responds differently from the typical relative, friend, or neighbor to the person who comes for help. This means she must be particularly alert to her own responses to what the client tells her. This leads to dilemmas and contradictions in the helping process that must be carefully thought through in order to maintain the integrity of the process. Some of these issues, dilemmas, and contradictions are discussed here.

3.1. Values and Ethics

Each of us has our own sense of what is right and wrong and our own preferences about what is important to us. Our ethics and values to a great extent determine who we are and how others see us. Furthermore, we tend to choose friends whose values and ethics are similar to our own. Some of us believe it is legitimate to use the loopholes in the law to reduce our taxes; others think this is cheating. Some of us prefer to spend our money on a house and car; others prefer to use our funds to give to charity or for political action. However, in most circumstances we are not in a position to influence others about what they think and what they value. Counselors, though, may be confronted with clients whose thinking about these and many other issues is quite different from their own.

What do they do about this? In most situations, if these values and ethics are unrelated to the problems the client brings to them, they do nothing. They have no right or obligation to attempt to change a client's values or ethics, even if they find them repulsive. If this makes it difficult for the counselor to work with the client, she must refer him to another counselor.

Sometimes these values are related to the client's problem. Then what? The use of argument or logic hardly ever works to change values, beliefs, or attitudes. Often, the only alternative for the counselor is to help the client see how these values are dysfunctional—how they may get him into trouble. Only when the client can make this connection is he likely to change. However, this can occur only if the counselor has established a relationship of trust with the client. One of the fundamentals of such trust is confidentiality.

We have noted that referral to another counselor may be necessary if the client's values and ethics are so different from those of the counselor that she finds it difficult to work with him. This may occur when there are major differences in culture and/or socioeconomic class between the client and the counselor. There are two other circumstances in which such a referral may be wise: when the counselor feels hostile to the client

and cannot figure out why and when the counselor believes that she does not have the expertise to be helpful to the client.

3.2. Confidentiality

Basic to a client's trust in the counselor is that she will not reveal to others the information given to her. Many helping professionals have what is known as privileged communication, which by law gives the counselor the right to hear almost anything the client says with the responsibility of not repeating it to others. A violation of confidentiality is likely to destroy the helping relationship and may even make it difficult for the client to seek help from any other counselor. Yet, if the client trusts the counselor, he may say things that seem outrageous to relatives and friends. He may tell the counselor that he hates his mother, that he wishes a child was dead, or that he is having an extramarital affair. Again, what makes the counselor special and helpful is that she reacts differently from friends and relatives. Above all, the information remains with the counselor except in the circumstances described next.

There are times when it is not only both ethical and legal to disclose information but also may be required. When a child or elderly person has been abused or neglected, the counselor is required to report this to the local protective service agency. When a client threatens another person with violence, the law requires that the counselor protect the potential victim by hospitalizing the possible perpetrator or informing the potential victim of the danger. Sometimes, it is a good idea for the counselor to discuss the case with a supervisor or to consult with another professional in deciding the legal, moral, and professional dimensions of her actions. Finally, the counselor may be required by the client's insurance company to provide certain information in order for the client to be reimbursed for her fees. However, in each of these instances, before confidentiality is violated, it is essential that the counselor get the client's permission, preferably in writing. This also provides her with the opportunity to discuss the issues related to the disclosure of information with the client and may open up new avenues for helping the client.

3.3. Trust

As implied previously, central to a successful or helpful counselor–client relationship is trust. However, this in itself is a dilemma. Often underlying the problem with which the client comes to the counselor is a lack of trust in those people with whom he has contact. Thus, one of the first and most important tasks of the counselor is to enable the client to learn to trust her.

What is the meaning of trust? In this context, trust means that the client believes that whatever the counselor does is in the client's best interest. Trust means that the client feels that no matter what he says it will not be used against him. Trust means that when the client is angry at the counselor, disagrees with the counselor's attitudes and values, or is ambivalent or indecisive, the counselor will not respond with anger or rejection, belittle, or badger him. Trust means that when the counselor confronts the client about his thoughts, feelings, and intentions, the client knows that the intent is to help him face up to the problems he is experiencing and that the counselor will provide support and encouragement as well. Trust means that the client understands that change is painful but that any pain the counselor inflicts on the client has as its purpose the reduction of the problem for which the client came into counseling.

However, if trust is so important, how does the counselor build it, especially since clients are unlikely to trust the counselor at the beginning of the counseling process? First, the counselor must be open and honest. The counselor must not evade the client's questions or respond with platitudes or try to give false hope. In her own way, she must convey to the client that although her attitudes and values may be different from his, each has a right to what each thinks and believes. The counselor respects the client for whom he is despite any differences there may be between them.

Second, the counselor encourages the client to lead and she is willing to follow, especially at the beginning of the counseling process. She allows the client to tell his "story" in his own way.

Third, the counselor encourages catharsis in order to relieve some of the client's anxiety, while at the same time she provides empathy and support. She listens with sensitivity and compassion, demonstrating by eye contact, body position, facial expression, and occasional words of reassurance that she is listening. However, this too can be a dilemma for the counselor, especially at the beginning of the counseling process. If too much anxiety is relieved through catharsis, the client is likely not to be motivated to return. On the other hand, if there is little or no relief of anxiety, the client is likely to think that the counseling process

is not helpful and again may not return. It is through experience that the counselor learns how to balance this issue in work with clients.

Thus, anxiety is a two-way process that has both negative and positive consequences. Most clients come to us because they have so much anxiety that they need to do something to reduce it. This is the motivation for change. On the other hand, many involuntary clients present a different type of dilemma. Many of them are not likely to be anxious. Many psychopaths and sociopaths do not experience anxiety in the same way most of us do. Without anxiety, there is no motivation for change, and without such motivation change is very difficult if not impossible. Thus, our job as counselors is to increase this type of client's anxiety level. One prominent way to achieve this is to point out the consequences of the client's behavior (e.g., loss of freedom by going to jail for dealing drugs or loss of life by overdosing on drugs). Nonetheless, this is a very difficult process.

Fourth, in order to establish trust, the counselor makes use of the client's priorities. That which the client sees as his most important or immediate problem must be dealt with. This is so even if the counselor believes that there are other problems that the client is experiencing that are more acute or may underlie the problem he brings to the interviewing session. The counselor must respect what the client believes is important.

A fifth method of gaining the client's trust is known as joining. This involves identifying with the client by following some of his behavior and speech patterns. Sometimes, it means mirroring the client's movements (leaning forward when the client does) or matching the client's timing (slowing down one's speech to be in harmony with the client's pace).

Sixth, if the counselor can provide immediate relief for one of the client's problems, trust is likely to occur. We have discussed how catharsis is one way of achieving this, but there are other ways as well. Sometimes, information can be very helpful. An explanation of unemployment or Social Security eligibility requirements can be very helpful with some clients. Explaining what is expected in college application materials to graduating high school students can relieve much anxiety. Referral to another agency or resource, such as a local tenants council or a store that sells reasonable clothing, can be seen as very helpful.

The "a-ha" response is another technique to establish trust. This occurs when the client arrives at a special understanding of his own situation that he did not have

before. It happens when the counselor is able to ask just the right question at the right time or make a crucial observation just when the client is ready to accept this about himself.

Finally, trust can and should be built by giving the client a sense of hope by assuring him that it is possible to change, to give up even the most painful feeling. However, for this to occur, the counselor must have the sense that she can be successful with this client. If the client does not have hope, the counselor must try to understand why and sometimes consult with a supervisor or colleague if this feeling persists. Without this sense of hope, it is highly unlikely that the counseling process can be helpful.

3.4. Diversity

Despite the counselor's good will, and her wish to be helpful to others, there are some clients who at first seem to reject the counselor's efforts and seem unwilling to trust her. These are people who are different from the counselor. They may be of a different race, may come from a different ethnic background or culture, or may be of a different gender or belong to a socioeconomic group different from that of the counselor's. They may have a difficult time understanding her, and she may have a difficult time understanding them. Her behavior may seem strange to them, and their behavior may seem strange to her. How can they trust someone who is so different from themselves?

Furthermore, if clients are from an oppressed group they may not only be anxious about the counselor's differences but also angry. Their experiences with "people like us" may have been very negative. The wheelchair-bound professional was denied employment despite his training and experience because of his disability. The black welfare mother was accused of child abuse when the teacher saw a bruise on the child even though it was caused by a fall. The gay or lesbian couple was denied health benefits because they were not legally married, although they have been living together for more than 20 years. The woman earns considerably less than her colleagues despite equal years of experience because of her gender. All of these people have reason to be angry, although the counselor is probably not the source of the anger.

How does the counselor deal with such differences? She must start by being a learner as well as a teacher. When the counselor does not understand the reasons for a client's behavior, it is her responsibility to figure out what is going on. She must not only teach the client

but also be open to learning from him. If the counselor is working primarily with a certain type of client—Puerto Ricans, the blind, or low-income women—it is useful for the counselor to read as much as she can about this group. However, this is not a substitute for asking the client about his behavior because not all clients fit the stereotypes described in the literature. By asking, the counselor not only learns about the client and his special differences but also shows respect for these differences.

However, what about the client who is angry? He may show that anger in a variety of ways; he may have a temper tantrum in the counselor's office, miss appointments, be unresponsive, give superficial answers to questions, or even lie. First, the counselor must remember that his anger is unlikely to be personally against her as an individual but rather who she represents. If she returns the anger, she will lose the client's trust. Rather, she needs to focus on the client's feelings, trying to determine the source of his rage. The counselor must let the client know that she understands that it may take him considerable time to trust her but that she is patient. She must try not to be defensive, such as by reassuring the client that she is not like other people. Although none of this may work for awhile, client anger has a way of wearing down, leaving the counselor a way of reaching the client.

3.5. Listening

The heart of the counseling process is the ability of the counselor to listen. She not only listens to the content of what the client is saying but also listens to how it is said. Is the client's speech a monotone or emotional, fast or slow, loud or quiet? Does it change depending on the topic being discussed? Does the tone of voice convey the image of a child, a seductive woman, an angry parent, or a sense of hopelessness or helplessness? Listening includes observation of the client as he speaks. Is there eye contact? Are there changes in facial expression or body position? All of these are clues about what is going on in the client's mind, and the client's thought processes provide the counselor with some understanding of the sources of the client's problem. The counselor must put all of this information together to assess the client's strengths and weaknesses and to plan with the client how she can best be of help to him. As she listens, the counselor must decide what response may be helpful to the client.

The counselor's responses are often different from those that might be expected from friends, neighbors, and relatives. The counselor raises questions that the client had not thought about before. The counselor makes statements that challenge the client's way of thinking. The counselor confronts the client's assumptions about himself and others. The counselor remains silent when the client expects a vigorous rebuttal. These are the responses that make the counseling relationship unique, and these are the behaviors that build trust between the client and counselor.

3.6. Silence

As unusual as it may seem, silence is an important part of the counseling process. Silence by the client has meaning that the counselor must try to understand. Silence by the counselor can be used to help the client in a variety of ways.

The client may be silent just because something he said or the counselor said has led him to think in new ways, to see new meaning or make connections he did not make before. He may be sorting out thoughts and feelings. This is a time when the counselor should also remain silent, as difficult as this may seem. When the client speaks again, a significant breakthrough in his thinking may occur.

The client may be silent because he is feeling deep emotion. The counselor lets the client experience this but not drown in it. She may ask, "What are you feeling right now?" The client may also be silent because he is confused. The counselor can interrupt the silence by indicating that what she said may be confusing and then try to clarify what has happened in the conversation.

The client may be silent because he is resisting the counselor's probing and/or rejecting the counselor's authority. There is tension in the air. The counselor's response might well be that both of them seem to be uncomfortable and maybe they should talk about what is causing the tension between them.

Thus, silence is not to be feared in the counseling process. Rather, it is to be understood and used by the counselor as a way of moving the helping process forward.

4. CONCLUSION

The counseling interview is a complex process that requires considerable training and experience by the counselor to be effective. Although the previous sections discussed the importance of listening and

silence by the counselor, it must not be forgotten that she must be in control of the interview to be effective. It is the counselor's ability to respond—and sometimes not respond—in ways that are not expected by the client that enables her to see the problem differently and deal with it more effectively. When successful, the counseling process can not only be of great aid to the client but also provide much satisfaction to the counselor.

See Also the Following Articles

Counseling and Culture ■ Ethics and Social Responsibility ■ Group Counseling ■ Measurement and Counseling ■ Trust

Further Reading

Benjamin, A. (1981). *The helping interview*. Boston: Houghton Mifflin.

Devore, W., & Schlesinger, E. G. (1991). *Ethnic-sensitive social work practice*. Columbus, OH: Merrill.

Donner, S., & Sessions, P. (1995). *Garrett's interviewing: Its principles and methods* (4th ed.). Milwaukee, WI: Families International.

Fine, S. F., & Glasser, P. H. (1996). *The first helping interview: Engaging the client and building trust*. Thousand Oaks, CA: Sage.

Lum, D. (1996). *Social work practice and people of color: A process stage approach* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Shulman, L. (1998). *Skills of helping individuals, groups, and communities*. Itasca, IL: Peacock.

Cross-Cultural Psychology, Overview

J. W. Berry

Queen's University, Kingston, Ontario, Canada

Harry C. Triandis

University of Illinois at Urbana–Champaign, Champaign, Illinois, USA

GLOSSARY

cross-cultural psychology The systematic study of behavior and experience as it occurs in different cultures, is influenced by culture, or results in changes in existing cultures.

cultural psychology The intensive study of psychological processes in one culture, which identifies how features of the culture affect these psychological processes.

indigenous psychology The systematic study of the culturally unique concepts and processes found in a given culture.

All human behavior is influenced by the culture in which a person develops. Thus, there can be no complete account of psychological phenomena without taking the cultural context into account. This claim applies not only to the findings of psychological researchers, but also to practitioners, who must consider the cultural settings of their applications as well as the cultural roots of the behavior.

1. DEFINITION OF CROSS-CULTURAL PSYCHOLOGY

According to Triandis and Lambert (1980), "Cross-cultural psychology is the systematic study of behavior and experience as it occurs in different cultures, is influenced by culture, or results in changes in existing cultures." And per Berry *et al.* (1997), "The field is diverse: some psychologists work intensively within one culture, some work comparatively across cultures, and some work with ethnic groups within culturally plural societies."

Some think of culture as an independent variable that influences psychological processes, and some emphasize that culture and psychological processes are mutually constituted, i.e., "make each other up." Berry argued that culture is both outside and inside the individual. It is independent of the individuals who arrive into a culture and become incorporated into a culture through the process of enculturation; but individuals also incorporate the culture through the same process, so that it also becomes an organismic variable, ready to be transmitted anew to other arrivals. In 2000, a special issue of the *Asian Journal of Social Psychology* discussed the

similarities and differences between indigenous, cultural, and cross-cultural psychology, which reflect these differences in point of view.

2. DEFINITION OF CULTURE

The conceptualization of culture is by no means a simple matter. One possible way to think about culture is that "culture is to society what memory is to individuals," as per Kluckhohn (1954). It includes what has worked in the experience of a society, so that is what was worth transmitting to future generations. It includes both objective elements, such as tools, bridges, chairs, and tables, and subjective elements such as norms, roles, laws, religions, and values.

Sperber used the analogy of an epidemic. A useful idea (e.g., how to make a tool) is adopted by more and more people and becomes an element of culture. Barkow, Cosmides, and Tooby (1992) distinguished three kinds of culture: metaculture, evoked culture, and epidemiological culture. They argued that "psychology underlies culture and society, and biological evolution underlies psychology" (p. 635). The biology that has been common to all humans, as a species distinguishable from other species, results in a metaculture that corresponds to panhuman mental contents and organization. Biology in different ecologies results in "evoked culture" (e.g., hot climate leads to light clothing), which reflects domain-specific mechanisms that are triggered by local circumstances and leads to within group similarities and between groups differences. What Sperber describes Barkow *et al.* call "epidemiological culture."

Elements of culture are shared standard operating procedures, unstated assumptions, tools, norms, values, habits about sampling the environment, and the like. Because perception and cognition depend on the information that is sampled from the environment and are fundamental psychological processes, this culturally influenced sampling of information is of particular interest to psychologists. Cultures develop conventions for sampling information from the environment and for how much to weigh the sampled elements. For example, people in hierarchical cultures are more likely to sample clues about hierarchy than clues about aesthetics. Triandis argued that people in individualist cultures, such as those of most cultures of North and Western Europe and North America, sample elements of the personal self (e.g., "I am busy, I am kind") with high probability. People from collectivist cultures, such as those of most cultures of Asia, Africa, and South America, tend to

sample mostly elements of the collective self (e.g., "my family thinks I am too busy, my co-workers think I am kind").

Just as individuals are incorporated into their primary culture through the process of enculturation, people may also become members of more than one culture through the process of acculturation, depending on life experiences.

2.1. Relationships with Other Perspectives

Cross-cultural psychology derives from both anthropology and psychology, taking various concepts, theories, and methods from each discipline. Because of this, the field is not unitary, but is composed of various perspectives. In our view, cross-cultural psychology combines two broad traditions. First, we use the comparative method (the "cross"), making assumptions about the comparability of behavior (see below) and employing research and statistical methods drawn primarily from individual psychology Second, the field is concerned with the unique sociocultural contexts in which people develop and act (the "cultural"), employing observational and narrative methods drawn primarily from cultural anthropology. The "cultural psychology" and "indigenous psychology" schools of thought emphasize this second perspective. However, in practice, both perspectives employ individual data, contextual data, and the comparative method in their work.

3. HISTORY OF CROSS-CULTURAL PSYCHOLOGY

The first person to write anything relevant about this topic was probably Herodotus, during the 5th century BC. There are three histories of the field that start with Herodotus, who had the insight that all humans are ethnocentric. This is necessarily an aspect of the human condition, because most humans are limited to knowing only their own culture and thus are bound to use it as the standard for comparing their culture with other cultures. It is unfortunate that this is necessarily the human condition, because it makes people think that their norms and values are universally valid, and any deviations from them is seen as not only wrong but also immoral. Much human conflict can be traced to ethnocentrism. It is only when humans have experienced several other cultures that they

become sufficiently sophisticated to see both the strengths and weaknesses of every culture and thus to become less ethnocentric.

Jahoda provided a detailed history of the field in which he discussed the work of Vico in the early 19th century, Wundt in the early 20th century, Tylor, Rivers, Boas, Luria, Bartlett, Bennedict, Vygotsky, and others. These writings were not central to the activities of most psychologists until the late 1980s and early 1990s, when many of the findings of cross-cultural psychology began challenging some of the theories of mainstream psychology.

The field became a distinct field of psychology in the early 1970s, when awareness of the need to stop being ethnocentric in research, and the need to develop measurements of psychological constructs that are both culturally sensitive and equivalent across cultures, became acute. Avoiding intellectual colonialism, such as the exploitation of data from other cultures without giving credit to local psychologists, resulted in an effort to develop a code of ethics for cross-cultural research. However, this code was not adopted by any association of psychologists because it was felt that it was overly strict and unrealistic and would inhibit research.

In the 1970s, most psychologists used an absolutist viewpoint. That is, it was assumed that all psychological discoveries were valid everywhere, and culture provided only minor modifications of these discoveries. In the 1980s, however, a number of researchers assumed more relativistic positions, in some cases even denying the psychic unity of humankind. Nowadays, a universalistic perspective is taken by many cross-cultural psychologists.

Journal publications with a cross-cultural psychology focus started in 1966. Leonard Doob, editor of the *Journal of Social Psychology,* and Germaine de Montmollin, editor of the *International Journal of Psychology*, announced that they would favor cross-cultural articles. The Human Relations Area Files started a journal that year that eventually became *Cross-Cultural Research*. It is now the official organ of the Society for Cross-Cultural Research. The *Cross-Cultural Newsletter* (first edited by Triandis) was launched in 1967. It eventually became the *Cross-Cultural Bulletin of the International Association for Cross-Cultural Psychology*. A directory of cross-cultural psychologists was first published by Berry in 1968. The *Journal of Cross-Cultural Psychology* (Walter Lonner, editor) started in 1970 and became the official organ of the International Association for Cross-Cultural Psychology. The Association pour la Recherche Interculturelle was established in 1984, and included French speaking cross-cultural psychologists. The Annual Review of Psychology introduced chapters reviewing cross-cultural work in 1973. The first edition of the *Handbook of Cross-Cultural Psychology* (Triandis, editor) appeared in six volumes in 1980–1981. The second edition (J.W. Berry, editor) appeared in three volumes in 1997.

3.1. Theory

The term absolutist was used to refer to a theoretical position that assumes that psychological findings are valid for all people everywhere; culture is thought to play little or no role in the development or expression of human behavior. Two other theoretical perspectives have been differentiated from absolutism. Relativism is essentially the opposite point of view: culture is so important in understanding human behavior that no psychological research or application can be done without knowing, and taking into account, the cultural context. An intermediate position (universalism) takes the position that human psychological processes are common, species-shared qualities; culture plays important variations on these basic processes during behavioral development and expression. In addition, absolutism tends to make comparisons freely; relativism generally avoids them; and universalism makes judicious comparisons (based on the underlying commonality) while taking care to understand the influence of culture. In a sense, the basic commonality makes comparisons possible (see discussion of equivalence, below), while the cultural variation in expression makes them interesting and worthwhile. In terms of the other perspectives or schools discussed previously, the cultural and indigenous approaches tend to be relativist, while the cross-cultural approach tends to be universalist. (But few psychologists who read this article are likely to be absolutist!)

4. METHODOLOGY

4.1. Levels of Analysis

Cross-cultural psychologists often work at different levels of analysis. For example, Hofstede has collected data from over 100,000 participants in more than 40 countries. By 2001, the data set was expanded to 50 countries. Responses to a sample of 40 and later 14 goals/value items were collected. He summed the responses to each goal/value item by the participants from each country. Then he factor-analyzed the 40 by

40 (or 14 by 14) matrix, based on 40 or 50 means. This is an ecological analysis and is very different from a within-culture analysis, in which the matrix of correlations is based on the number of participants from that culture.

Readers may need an illustration of the fact that results at the cultural (or group) and individual levels of analysis can be different. Lincoln and Zeitz studied 500 employees divided among 20 social service agencies. They found that supervisory duties were positively correlated with professional qualifications across individuals but negatively correlated across agencies. The latter correlation was due to the fact that the members of the more professional agencies did not need much supervision.

4.2. Comparability and Equivalence

As was noted above, there are different viewpoints regarding whether comparisons are legitimate: relativists tend to object to them, absolutists tend to make them without restriction, and universalists employ them with caution, once certain conditions are met.

All acts of comparison (whether they be of cultures, institutions, or behaviors) require some common dimension along which phenomena can be judged. Apples and oranges can indeed be compared: on sugar content, juiciness, or price. The presence of such commonalities is a prerequisite for making a comparison; however, it is not enough. Cross-cultural psychologists seek to establish the comparability of two (or more) phenomena, using a number of ideas and techniques. One of these is equivalence, in which conceptual meaning and psychometric properties can be demonstrated to be sufficiently similar between two sets of data to be compared. It is no longer accepted practice to simply use an instrument, obtain a score, and apply a test of significance to the difference. While this has been known for a long time, it is still ignored by some writers who submit papers to the editors of psychological journals.

4.3. Emics and Etics

One approach to better controlling cross-cultural comparison is with the use of the concepts of emics and etics. All cultures categorize, but the categories that are used are culture specific. For instance, our eye is capable of discriminating 7,500,000 million colors, but we get along with a couple of dozen color names. Through the use of Munsell color chips, which provide an objective domain to which we can compare the way people in different cultures categorize colors, we can determine that the categories are not identical. The stimuli at the core of major color categories such as "red" are the same, but the periphery is not. So, even when we deal with something as simple as the identification of colors we must take cultural specificity into account. All categories are culture specific. We call these emic, from Pike's use of phonemics and phonetics. But while most categories are emic, there are elements that are universal, which we call etic. In short, if we stick only to the core of each color range (the etic part) we can use the word "red" and its translations across cultures with reasonable equivalence. But there are regions of the category that are not equivalent across cultures.

Furthermore, there are languages that have very few color names. Some languages have only two words that are relevant to color that are more or less equivalent to white and black. When a language has three color names, it adds a word for red. If it has four or five names, then it also has words that are more or less equivalent to yellow and green. Languages that have progressively richer color vocabularies have a word for blue, then for brown, finally for purple, pink, orange, and gray.

People who do not distinguish between green and blue have difficulties in discussing differences in color between grass and the sky. If we study this domain and assume that they see color the way we do, we are likely to obtain distorted results. To avoid this problem, when we do a study we should talk to many people from each language to see how they cut the pie of experience. This would identify their emics for dealing with the particular topic. When we do that, some of their categories may also appear in other languages. Then, they may in fact be etics.

When we construct instruments such as tests or interview questions, or scales, we need to incorporate both emics and etics. The emics allows us to measure a phenomenon with cultural sensitivity, "the way the natives see it." The etics allows us to compare the cultures. For example, a scale of social distance was constructed for Greeks and Americans in 1962. The scale values were standardized separately for each culture, and were made equal to zero for "I would marry this person" and 100 for "As soon as I have a chance I am going to kill this person." The item "I would accept this person as a roommate" made sense only in the United States and had a scale value of 29.5. The item "I would accept this person in my parea" was a Greek emic, and had a value of 31.1. A parea is an intimate group of friends that meets almost daily to enjoy things together. "I would rent a room from this person" had a scale value

of 57.5 in the United States and 42.8 in Greece, suggesting that in Greece it was a more intimate activity. In a culture such as India, where touching is a very important issue, one would use emic items such as "I would allow this person to touch my earthenware." In a culture that has residential segregation, an item such as "I would allow this person to live in my neighborhood" would be a useful item to measure social distance.

We also need to consider the fact that often a construct used in the West has very different meanings in other parts of the world. For example, when people in the West use the concept of "intelligence," they include the ideas of speed, successful adaptation, and effective completion of tasks. In parts of Africa, people think that to be intelligent means to be slow, sure, and not make mistakes. In other parts it includes the ideas "knows our traditions" and "does what the elders expect." Among the Cree, a Native American tribe, "lives like a white" is conceptually close to "stupid."

4.4. Translation

In the construction of instruments that can be used equivalently across cultures, researchers need to use translation. This is a complex process.

5. DIMENSIONS OF CULTURAL VARIATION

As has just been noted, comparisons require common dimensions on which they can be made. Beyond this simple dictum is a complex reality: there are many possible dimensions that could be employed in comparing cultures, institutions, and behaviors.

When we think of the relationship between culture and psychology, it is useful to use a metaphor from geography. We can discuss continents by describing mountain ranges, rivers, bays, etc. We can discuss countries by describing similar features, cities, villages, roads, etc. We can describe cities by describing roads, buildings, etc. We can describe villages by specifying the location of every building in relation to every other building. We can describe neighborhoods by discussing the economic, social relationships, religions, and political organizations of the residents, etc. Finally, we can discuss households by giving the biography of each member of the household, etc.

Note that as we move to lower levels of abstraction, we need more dimensions, details, specifications.

Psychologists deal with phenomena that are based on the common biology of humans, as it adapts to different environments (continents). Social psychologists also examine different environments (ecologies), for instance, social relationships in which people influence each other a great deal, as one finds in agricultural societies, or relatively little, as one finds among hunters and gatherers. Then, different behaviors can be identified. For example, there is more conformity in agricultural than in hunting environments. The societies that are agricultural have more inequality than the societies that engage in hunting and gathering.

Different disciplines focus on different levels of analysis. For example, literary productions describe events at the level of households. Anthropologists describe neighborhoods. Sociologists tend to describe cities and villages. We believe that all this work is valuable. All the information is necessary for a balanced understanding of the culture and psychology relationships.

In this section we focus on some of the dimensions of cultural variation that correspond to the level of continents. The data are obtained from different countries, though this is a very rough procedure because each country contains hundreds of cultures. Yet because regions of the world (e.g., Europe versus East Asia) have something in common, we can discuss the ways in which they are similar and different. Other work that we have no space to discuss is done at the other levels of analysis and is also valuable. The dimensions of cultural comparison are also called "cultural syndromes."

5.1. Complexity

We mentioned previously the contrast between hunters and gatherers and agricultural societies. Agricultural societies are more complex in the sense that they have more types of roles than hunting and gathering societies. As we move to modern information societies we find even more complexity, for instance, many more roles than in more simple cultures. If we consider all forms of specialization, e.g., the difference between medical specializations, in complex societies there are more than a quarter of a million roles.

5.2. Tightness and Looseness

Societies differ in the number of norms concerning correct behavior. Some have many norms, e.g., how to bow and how to smile, and others have few norms. Furthermore, some cultures punish severely those who deviate from norms whereas others are lenient. In tight

cultures, there are many norms, and those who do not follow the norms are criticized or even killed. In loose societies, there are not many norms, and people who do not behave according to the norms are not punished. People are very likely to say "It does not matter" when they see a deviation from a norm. An example of a tight culture was the Taliban in Afghanistan. They killed people in large numbers for relatively minor deviations from arbitrary norms that they imposed on the basis of supposedly religious guidance. Some of these norms, such as not listening to music and not watching Western television, had nothing to do with religion. An example of a loose society, on the other hand, is Thailand, where people tend to smile even when employees do not show up for work as expected. The imposition on norms depends of the domain or the situation. For example, in the United States there is tightness concerning how to deal with banks, your checkbook, loans, mortgages, and the like. But now there is considerable looseness concerning who you can live with. Of course, there are also major individual differences. Some people are compulsive about their own and others' behavior, while others are very lenient (e.g., any behavior is okay; it's none of my business to criticize this behavior).

Hofstede identified a dimension he called uncertainty avoidance, which is related to tightness. Countries high on this dimension were Greece, Portugal, and Guatemala; low were Denmark, Jamaica, and Singapore. Subjective well-being was negatively correlated with this dimension, presumably because people in tight cultures are anxious about being criticized or punished for behavior that is not perfectly "correct."

5.3. Collectivism and Individualism

Societies that are both simple and tight tend to be collectivist. Societies that are both complex and loose tend to be individualist. Collectivism is a cultural pattern found especially in East Asia, Latin America, and Africa. It is usually contrasted with individualism found in the West, e.g., in Western and Northern Europe, the United States, Canada, Australia, and New Zealand. Note, however, that any typology is an oversimplification; societies are not purely individualist or collectivist but some mixture of the two. The construct has been widely used in the study of cultural differences. Unfortunately, however, most of the research compared samples from North America versus East Asia. Thus, in the description of the attributes of the construct and its antecedents and consequents we rely too heavily on the differences between the United States and Canada on

the one hand and China, Korea, and Japan on the other hand. We do not know if the information we present below applies equally to Africa and Latin America, though we suspect that much of it does. Much more research is needed on this topic.

Among the most important characteristics of the collectivist cultural pattern are the following:

1. Individuals define themselves as aspects of a collective, interdependent with some ingroup, such as one's family, tribe, co-workers, or nation, or a religious, political, ideological, economic, or aesthetic group.
2. They give priority to the goals of that collective rather than to their personal ones.
3. Their behavior is determined more often by the norms, roles, and goals of their collective than by their personal attitudes, perceived rights, or likes and dislikes.
4. They stay in relationships even when the costs of staying exceed the advantages of remaining.

These are the defining attributes of collectivism. There are as many kinds of collectivism as there are collectivist cultures. To distinguish among collectivist cultures additional attributes are necessary.

Triandis summarized the literature that used that construct in social psychology, and Kagitcibasi provided a critical evaluation of the construct. Hofstede found that the United States, Australia, and Britain were highest in individualism and Panama, Ecuador, and Guatemala were highest in collectivism. Most Western countries are individualist and most East Asian countries are collectivist. However, gross national product per capita is correlated with individualism; thus, wealthy East Asian countries such as Japan are not as collectivist as one would expect from their geographic location.

Within any society there are individuals who behave like persons in collectivist cultures, called allocentrics. They contrast with those who behave like persons in individualist cultures, who are called idiocentrics. There are both allocentrics and idiocentrics in every society, but their distributions are different, with more allocentrics found in collectivist cultures.

Collectivism is maximal in relatively homogeneous societies, such as theocracies and monasteries, while individualism is maximal in heterogeneous societies that are very affluent. Thus, there will be few idiocentrics in monasteries and few allocentrics among Hollywood stars.

All individuals have access to cognitive systems that include both allocentric and idiocentric cognitions, but they sample them with different probabilities, depending on the situation. For example, if the ingroup is under

attack, most individuals become allocentric. In the company of other allocentrics, the norms for allocentric behavior become salient, and individuals are more likely to sample allocentric cognitions. Some situations provide very clear norms about appropriate behavior (e.g., in a house of worship), while other situations do not (e.g., at a party). Individuals will be more allocentric in the former than in the latter situations. When the ingroup can supervise an individual's behavior, norms are more likely to be observed, and the individual will be more allocentric. In one study, allocentrics and idiocentrics were randomly assigned to simulated organizations that were individualist or collectivist, and the degree of cooperation exhibited by the individuals was studied. It was found that cooperation was high when allocentrics were in a collectivist organization. In the other three cells of the experiment, there was very little cooperation. In short, both the personality and the situation were required to predict the behavior.

Collectivism can appear in all or none of the domains of social life. For example, it can be found in politics, religion, aesthetics, social life, economics, or philosophy, as was the case in China during the Mao period, or in none of these domains, as among Hollywood stars.

5.3.1. Antecedents of Collectivism

Some of the antecedents of collectivism are listed in Table I and are discussed below:

Homogeneity of the collective. If people disagree about the norms of proper behavior, or the goals that people should have, it is difficult for people to behave according to the norms of the group.

TABLE I
Some Antecedents of Collectivism

- Homogeneity of the collective
- Low exposure to other cultures
- Interdependence is adaptive to ecology
- Population density
- Social stability
- Low economic status
- Large families
- One large ingroup rather than many ingroups
- Salient common fate
- Collective mobility is advantageous
- Strict religious upbringing
- Little exposure to modern mass media

Low exposure to other cultures. People who know only one culture tend to be maximally ethnocentric, authoritarian, and submissive to ingroup authorities. Those who are more educated, traveled, and have lived with more than one cultural group develop idiocentric tendencies.

Interdependence is adaptive to ecology. People are more interdependent in agricultural societies than in information societies. For example, when the goal is to complete large projects such as irrigation canals or defensive walls, collectivism is more likely. In societies where people are financially interdependent, collectivism is high. People who can do their job when they are alone are more likely to be idiocentric.

Population density. In dense social environments, many rules that are designed to reduce conflict and ensure the smooth functioning of the group develop.

Social stability. When the collective is stable, it is more likely to develop agreements about norms, and to make sure that the norms are observed. There is evidence that the older members of all societies are more allocentric than the younger members.

Low economic status. The lower social classes tend to be more conforming to social norms than members of the upper classes. When resources are limited, one often depends on group members for assistance, especially in emergencies. These factors increase collectivism. On the other hand, in all cultures, those in positions of leadership tend to be idiocentric.

Large families. In large families it is not practical to allow each child to follow idiosyncratic schedules or to have much privacy. Many rules are enforced, and that creates collectivism.

One large ingroup rather than many ingroups. Those who only have one ingroup can channel all their energy into that group. Also, they cannot afford to develop poor relationships with members of that group, so they are more likely to observe its norms.

Salient common fate. Common fate with members of the ingroup (e.g., when the ingroup is under attack) increases collectivism. Time pressure for decisions has similar consequences.

Collective mobility is advantageous. If individual upward social mobility is not possible, then collective mobility may be used. Thus, individuals invest their energy in promoting the status of their ingroup.

Strict religious upbringing. Most religions require observance of a large number of norms and threaten to punish those who ignore these norms. That increases collectivism.

Little exposure to modern mass media. U.S.-made television is widely available throughout the world.

Content analyses show that the themes used are highly individualistic (e.g., emphasis on pleasure, doing what the individual wishes to do even if that is inconsistent with the wishes of authorities). Countries where people have little exposure to Western mass media are more collectivist.

5.3.2. Consequences Of Collectivism for Cognition, Emotion, and Motivation

Marcus and Kitayama have reviewed much evidence showing the influence of collectivism on cognition, emotion, and motivation (see Table II).

Thinking. Nisbett, Peng, Choi, and Norenzayan examined a broad range of literature concerning thought patterns in ancient Greece and China, two cultures that correspond to the individualism–collectivism contrast. They argued that the Greeks thought analytically and the Chinese holistically. The Greek traditions influenced the West, which now uses logic, rules, and categories and sees the world as a collection of discrete objects, while the East sees the world as a collection of overlapping and interpenetrating substances. They reviewed much empirical work that contrasted Western and Eastern cultures, and found that the West focuses on objects and their attributes, while the East focuses on the context of objects and the relationships of objects and events to each other. Western thought is linear, concerned with abstract analysis, and pays much attention to contradictions. Eastern thought is dialectical, experience-based, and pays little attention to contradiction. The West is more field independent; the East more field dependent. Explanations of events in the West focus on the attributes of the object, e.g., the attitudes, beliefs, and values of individuals as determinants of behavior; in the East, the focus is on the context of the object, e.g., group memberships, norms, and social pressures that influence the behavior of individuals. Numerous other cognitive patterns contrast these two kinds of cultures and show not only quantitative but also qualitative differences in thinking.

Attributions. Allocentrics make external attributions in explaining events, whereas idiocentrics show the opposite pattern. Thus, the idiocentrics make the fundamental attribution error (i.e., seeing internal rather than external causes for the behavior of other people, while the actors themselves see external causes as explanations of their own behavior more often than internal causes) more frequently than do allocentrics. Also, allocentrics attribute their successes to the help they have received from others and their failures to personal shortcomings (e.g., I did not try hard enough). Allocentrics tend to think that social environments are stable and individuals are easy to change so that they can fit into the environment; idiocentrics think that individuals are stable and social environments are easy to change.

Emotions. Members of collectivist cultures get angry when something unpleasant happens to ingroup members more often than when something unpleasant happens to themselves. They do not express negative emotions as often as do members of individualist cultures.

Description of self. When asked to complete 20 statements that begin with "I am . . ." members of collectivist cultures provide sentence completions that have more social content (e.g., I am a brother, I am a member of the Communist party) than do members of individualist cultures.

Socialization. In collectivist cultures, children are socialized to be obedient, be reliable, follow traditions, do their duties, sacrifice for the ingroup, and be conforming. An important value is self-control.

Low well-being. While members of all cultures are satisfied with their life, members of collectivist cultures are less satisfied than members of individualist cultures. Well-being tends to be correlated with individualism. This relationship persists even when statistically controlling for income.

TABLE
Consequences of Collectivism

- Thinking
- Attributions
- Emotions
- Describe the self by using roles, groups
- Consequences of socialization
- Low well-being
- Social behavior
- Receive much social support if they conform
- Have poor skills for entering new groups
- Are easy to mobilize for confrontations with outgroups
- Have the ideal of fitting-in rather than shining
- Show little social loafing
- Recreation occurs in stable groups
- Communication
- Conflict resolution
- Conformity
- Apologize more for transgressions of the ingroup

Social behavior. Collectivists are extremely concerned with "saving face," and even attempt to save the face of the persons with whom they are interacting. Preserving harmony within the ingroup and being pleasant in relationships with ingroup members are often important values. On the other hand, collectivists do not care much about their relationships with outgroups. They usually behave quite differently toward ingroup versus outgroup members, showing sacrifice and extreme cooperation with ingroup members and suspicion and hostility toward outgroup members.

Social support. Allocentrics report that they receive more social support and a better quality of social support than do idiocentrics. However, those who do not conform to ingroup norms are rejected, and may even be killed.

Poor skills for entering new groups. Members of collectivist cultures do not have good skills for entering new groups, because they are not socialized to do this. They are comfortable as members of their ingroups, thus they do not need to enter other groups.

Easy to mobilize for confrontations with outgroups. When ingroups are in conflict it is easier to arouse members of collectivist cultures to fight for the ingroup than it is to do so with members of individualist cultures. Ethnic cleansing is more often found in collectivist than in individualist cultures.

Have the ideal of fitting-in rather than shining. Allocentrics prefer to fit in and be like most others in their ingroup rather than shining and sticking out. Idiocentrics tend to be high in self-enhancement (i.e., report that they are better than most other people on most desirable attributes). Allocentrics tend to be modest when they compare themselves to others.

Show little social loafing. Allocentrics show little evidence of social loafing (i.e., the tendency to produce less when working in a group versus when working alone, if their output is not identifiable). They usually produce as much as possible when they work with ingroup members, but they do show social loafing if their co-workers are outgroup members.

Recreation. Members of collectivist cultures seek recreation in groups to a greater extent than members of individualist cultures (e.g., bowling rather than skiing).

Communication. Members of collectivist cultures rely on the context (e.g., distance between and position of bodies, eye contact, gestures, tone of voice) more than on the content of communications. For example, rather than saying "no," they may serve incongruous foods in order to put across their negative message without causing the other person to lose face. When they describe a person they are likely to use context, e.g., "He is intelligent in the marketplace," rather than decontextualized statements, such as "He is intelligent." Silence is used more in collectivist than in individualist cultures. Compliments are not used as frequently as in individualist cultures. The languages of collectivist cultures do not require the use of "I" or "you." Collectivists use more action verbs (e.g., he asked for help) rather than state verbs (e.g., he is helpful), which suggests that they use communications that make greater use of context, are more concrete, and place less emphasis on the internal attributes of the person.

Conflict resolution. Members of collectivist cultures avoid conflict with ingroup members, and in the case of disagreement they prefer silence to argument. They prefer mediation to confrontation.

Conformity. Members of collectivist cultures show more conformity than members of individualist cultures in experiments that measure conformity.

Apologize more for transgressions of the ingroup. Members of collectivist cultures apologize more when an ingroup member commits a crime than is common in individualist cultures.

5.3.3. Evaluation of the Construct

While people in collectivist countries indicate that they are less happy than people in individualist countries, individualism has been found to be associated with poor marital adjustment, thus high divorce rates, and more delinquency, crime, drug abuse, suicide, experimenting with early sex, and stress.

Naroll has reviewed much evidence that indicates that tight social groups are needed for social control and the good life. Diener and Suh suggest that "real happiness" depends on having many small positive experiences rather than having a few very intensive positive experiences. Thus, while people in collectivist countries may feel unable to do "their own thing" and thus report low levels of happiness at the moment they are asked if they are happy by the researcher, they may be happier in the long run. Thus, each cultural pattern has both positive and negative aspects.

5.4. Power Distance

Hofstede identified a dimension of cultural variation he called power distance. Cultures high on this dimension have members who see a large distance between those

who have power and those who do not. Thus, for instance, people in such cultures fear disagreement with their superiors. Countries high on this dimension were Malaysia, Guatemala, and Panama. Countries low were Denmark, Israel, and Austria. This dimension was highly correlated with collectivism. It was also correlated with the level of corruption of the country.

5.5. Sex Differentiation

Hofstede has identified a dimension he calls masculinity–femininity. In feminine cultures, members of the culture attach more importance to relationships, to helping others, and to the physical environment than people in masculine cultures. In masculine cultures, people emphasize careers and money. The goals of men and women are more differentiated in masculine than in feminine cultures. National samples that emphasized male goals were found in Japan, Austria, and Venezuela, whereas the Scandinavian countries and the Netherlands emphasized female goals. In the countries that were high in masculinity, men and women gave very different responses to Hofstede's value questions, while in the countries that emphasized female goals, the answers obtained from men and women were the same. Similar differences were observed across occupations, with engineers giving masculine and secretaries giving feminine answers. The consequences of this cultural difference included such matters as people in masculine countries wanting brilliant teachers versus people in feminine countries wanting friendly teachers. Differences in politics were also identified. In masculine countries people supported tough policies toward poor people and immigrants. Economic growth was given priority over preservation of the environment in masculine countries. In feminine countries, compassionate policies toward the weak and the environment were given high priorities.

5.6. Dealing with Time

There are cultural differences in the way time is viewed and in the speed of life. People in countries that are rich, individualistic, and in cold climates move fast and arrive for appointments at the correct time. Those in poor, collectivist countries located in hot climates tend to move slowly and arrive for appointments late. For example, because in collectivist cultures interpersonal relations are very important, if a person who is on her way to an appointment meets a friend, the chances are that she will spend much time with the friend and be very late for the appointment. But it will not be

necessary for her to apologize for being late, as most likely the other people at the appointment will also arrive late.

In all countries, cities are faster than rural environments. Levine devised a number of methods that objectively measured speed; for example, how fast do people walk in the street, how long does it take to mail a letter, how accurate are the clocks in public places. These measures were intercorrelated, showing concurrent validity. He took several samples on each measure in 31 countries. Switzerland, Ireland, and Germany were the fastest and Brazil, Indonesia, and Mexico the slowest countries.

Hofstede introduced a measure of long- versus short-term orientation. A long-term orientation was found in East Asia, among countries with a Confucian tradition, while a short-term orientation was found in the Philippines, Nigeria, and Pakistan.

5.7. Missing Topics

The relationship between culture and psychology is so extensive and complex that in this article we were able to present only an introduction. Cultural differences in trust, although not discussed here, are especially important because in countries that are low in trust it is difficult to have economic development. Low trust is frequently found in collectivist cultures where outgroups are often seen as enemies. Because most people belong to outgroups, people in low-trust countries feel surrounded by enemies. A study by Bond, Leung, and 60 other researchers identified a factor that is related to distrust. Countries such as Pakistan and Thailand were high on this factor whereas most of the European countries were low. It is not clear at this time if economic development results in high trust or if high trust leads to economic development. It is likely that the relationship is reciprocal.

6. EVALUATION OF FINDINGS AND APPLICATIONS

Approximately 20 years ago, the World Health Organization (WHO) approached the International Association for Cross-Cultural Psychology with what appeared to be a rather simple question: "What do you know, as cross-cultural psychologists, that can help the WHO to achieve its goal of Health for All by 2000?" A number of meetings were held to discuss, and to try to answer, this question. The result was a volume of topical appraisals and an evaluative overview of them. Dasen

et al. concluded that there was considerable research-based knowledge, but little that was readily applicable. Two reasons were advanced to explain this discrepancy. First, most knowledge acquisition is driven by a research orientation more than (even rather than) a view to eventual application: at the end of a project, there is often the question, "What do we do with all this?" Second, most research, even research that is explicitly cross-cultural, is rooted in one dominant psychological perspective (e.g., WASP, Western academic scientific psychology) that smacks of ethnocentrism, even if that is inadvertent.

A later approach was made to a broader range of questions that went beyond health to other domains of human activity. Once again we concluded that there was great potential but little actual achievement: a substantial knowledge base had not yet been applied. More recently, a number of authors addressed this question for a limited number of behavioral domains, and found that while progress had again been made, the gap between knowledge and practice remained wide. This state of affairs is not unique to cross-cultural psychology, but it is made more salient by the urgent need to use psychological knowledge in many parts of the world. Critical projects are needed in societies other than those in which the knowledge was generated.

One way to assist in the process of application is to present the knowledge base for findings that may qualify as "universals," and hence that may be applicable across cultures. In such cases, the problem of cultural irrelevance or incompatibility may be reduced, even though the science–application gap may remain. Following the sequence in this article, we can first make use of the knowledge provided by our cognate disciplines of cultural anthropology, linguistics, and sociology. These disciplines have established universals in sociocultural systems and institutions, ones that can be found in all societies. Some of these have been portrayed in Section 5. What we take from these is not that the variation hinders application, but that the common dimensions provide a basis for the cross-cultural use of various research findings.

From these other disciplines, we have also adopted theoretical perspectives (especially relativism and universalism) and concepts that provide foundations for our methodological tools (emics, etics, equivalence); other applications include cultural similarities and differences in organizational behavior. Finally, an important application is training people to learn how to interact successfully with those from other cultures.

See Also the Following Articles

Cognition and Culture ■ Cultural Psychology ■ Cultural Syndromes ■ Gender and Culture ■ Indigenous Psychologies ■ Intelligence and Culture ■ Intergroup Relations and Culture ■ Leadership and Culture ■ Perception and Culture ■ Power, Authority, and Leadership

Further Reading

Aberle, D. F., Cohen, A. K., Davis, A., Levy, M., & Sutton, F. X. (1950). Functional prerequisites of society. *Ethics*, *60*, 100–111.

Barkow, G., Cosmides, L., & Tooby, J. (Eds.) (1992). *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Berry, J. W. (1969). On cross-cultural comparability. *International Journal of Psychology*, *4*, 119–128.

Berry, J. W. (1975). *Applied cross-cultural psychology*. Lise: Sets & Zeitlinger.

Berry, J. W. (1976). *Human ecology and cognitive style: Comparative studies in cultural and psychological adaptation.* Beverly Hills, CA: Sage.

Berry, J. W. (2000). Cross-cultural psychology: A symbiosis of cultural and comparative approaches. *Asian Journal of Social Psychology*, *3*, 197–205.

Berry, J. W., Poortinga, Y. H., & Pandey, J. (1997). *Handbook of cross-cultural psychology* (Vol. 1, 3rd ed.) Boston, MA: Allyn & Bacon.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications*. New York: Cambridge University.

Brislin, R. (Ed.) (1990). *Applied cross-cultural psychology*. Newbury Park: Sage.

Dasen, P. R., Berry, J. W., & Sartorius, N. (Eds.) (1988). *Health and cross-culture psychology: Towards applications.* Newbury Park: Sage.

Diener, E., & Suh, E. M. (Eds.) (2000). *Subjective well-being across cultures.* Cambridge, MA: MIT Press.

Diener, E., Diener, M., & Diener, C. (1995). Factors predicting the subjective well-being of nations. *Journal of Personality and Social Psychology*, *69*, 851–864.

Frijda, N., & Jahoda, G. (1966). On the score and methods of cross-cultural research. *International Journal of Psychology*, *1*, 110–127.

Greenfield, P. M. (1997). You can't take it with you. *American Psychologist*, *52*, 1115–1124.

Hofstede, G. (2001). *Culture's consequences* (2nd ed.). Thousand Oaks, CA: Sage.

Inglehart, R., & Baker, W. (2000). Modernization cultural change and the persistence of traditional values. *American Sociological Review*, *65*, 19–51.

Jahoda, G. (1992). *Crossroads between culture and mind.* Cambridge, MA: Harvard University Press.

Jahoda, G., & Krewer, B. (1997). History of cross cultural and cultural psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross cultural psychology* (2nd ed., pp. 1–42). Boston, MA: Allyn & Bacon.

Kagitcibasi, C. (1997). Individualism and collectivism. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology* (2nd ed., pp. 1–50). Boston, MA: Allyn & Bacon.

Klineberg, O. (1980). Historical perspectives: Cross-cultural psychology before 1960. In H. C. Triandis, & W. W. Lambert (Eds.), *Handbook of cross-cultural psychology* (Vol. 1, pp. 1–14). Boston, MA: Allyn & Bacon.

Kluckhohn, C. (1954). Culture and behavior. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 921–976). Cambridge, MA: Addison-Wesley.

Lenski, G. E. (1966). *Power and privilege: A theory of social stratification*. New York: McGraw-Hill.

Levine, R. (1997). *A geography of time*. New York: Harper Collins.

Lincoln, J. R., & Zeitz, G. (1980). Organizational properties from aggregate data: Separating individual and structural effects. *American Sociological Review*, *45*, 391–408.

Markus, H., & Kitayama, S. (1991). Culture and self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Naroll, R. (1983). *The moral order*. Beverly Hills, CA: Sage.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*, 291–310.

Murdock, G. (1967). *Ethnographic atlas*. Pittsburgh, PA: University of Pittsburgh Press.

Shweder, R. A. (1990). Cultural psychology: What is it? In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 1–43). Cambridge, MA: Cambridge University Press.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford, UK: Blackwell.

Tapp, J. L., Kelman, H. C., Triandis, H. C., Wrightsman, L., & Coelho, G. (1974). Continuing concerns in cross-cultural ethics: A report. *International Journal of Psychology*, *9*, 231–249.

Triandis, H. C. (1964). Cultural influences upon cognitive processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 1–48). New York: Academic Press.

Triandis, H. C. (1989). Self and social behavior in differing cultural contexts. *Psychological Review*, *96*, 506–520.

Triandis, H. C. (1994). *Culture and social behavior*. New York: McGraw-Hill.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

Triandis, H. C., & Lambert, W. W. (Eds.) (1980). *Handbook of cross-cultural psychology: Perspectives* (Vol. 1). Rockleigh, NJ: Allyn and Bacon.

Witkin, H. A., & Berry, J. W. (1975). Psychological differentiation in cross-cultural perspective. *Journal of Cross-Cultural Psychology*, *6*, 4–87.

van de Vijver, F., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.

Cultural Complexity

Mark W. Allen

California State Polytechnic University, Pomona, California, USA

James H. Liu

Victoria University of Wellington, Wellington, New Zealand

GLOSSARY

individualism–collectivism Continuous dimension of cultural variation emphasizing an orientation toward personal goals, independent self-construal, and individual autonomy, at the one end, and group goals, interdependent self-construal, and emphasis on long-standing close relationships within in-groups, at the other.

looseness–tightness Continuous dimension of cultural variation describing the degree to which a culture endorses multiple and pluralistic standards of behavior, at the one end, and uniform, socially prescribed behaviors, at the other; loose cultures tolerate deviance from norms, whereas tight cultures emphasize sanctions for violating norms.

multilineal cultural evolution Dominant perspective on the evolution of cultural complexity; although it emphasizes some universal patterns in culture change from family level to state societies, it acknowledges multiple causal variables and trajectories in cultural evolution.

power distance Continuous dimension of cultural variation describing how relationships between high-ranking and low-ranking persons are conducted, with a high degree of deference to superiors and social distance between superiors and subordinates, at the one end, and egalitarianism, at the other.

social power The means by which a social actor can influence, motivate, or control other individuals; it can be developed through interwoven networks of economic, military, ideological, and political sources of power and control.

standard cross-cultural sample (SCCS) A coded sample of worldwide cultures compiled by anthropologists for the purposes of cross-cultural studies.

unilineal cultural evolution Popular theoretical perspective during the late 19th century that cultures evolve along a single lineal dimension through set stages, usually based on technological or economic development; it has been long discredited by cultural anthropology and archaeology as simplistic and ethnocentric.

Cultural complexity entails the presence of a multitude of social networks, social groups, hierarchies, power structures, roles, positions, and/or divisions. Whereas kinship is the primary organizational concept in simpler societies, complex societies have multiple levels of crosscutting ties, affiliations, responsibilities, membership, values, norms, and rights. Individuals pursue their own particular goals through social networks, and no two people have the exact same network or objectives. Perhaps the key difference between simple societies and complex societies is the presence of social power, that is, the means by which human actors can manipulate others in the pursuit of social goals.

1. CULTURAL COMPLEXITY: A RECENT PHENOMENON

Cultural complexity is a recent phenomenon of human society. Paleoanthropology and the fossil record reveal a record of at least 5 million to 6 million years of hominid evolution. Hominids with large brain cases, the genus *Homo,* date back to circa 2 million years ago. Archaic *Homo sapiens,* including the Neanderthals, stretch back to perhaps 400,000 years ago. Our own species, *Homo sapiens,* can trace its origin to Africa some 150,000 to 200,000 years ago.

Yet even after the appearance of physically modern humans (with brains apparently as large and complex as our own), human culture and society remained relatively simple for tens of thousands of years. Despite having the same mental "hardware" as we do today, humans showed no flashes of art, complex organization, ethnicity, or other indications of significant change until the Upper Paleolithic period in the Old World circa 40,000 years ago. Even so, this early fluorescence of art, religion, technology, ethnicity, identity, and other human cultural hallmarks was still characterized by small and mobile populations, the lack of institutionalized leadership, extremely limited social power, and fairly simple social networks.

It is not until the first appearance of dense population, food surpluses, and settled villages roughly 10,000 years ago (the Neolithic period in the Old World and somewhat later during the Formative period of the New World) that cultural complexity can first be discerned in the human record. Since that point, there has been a drastic trend toward fewer and more complex societies. Eventually, chiefdoms, archaic states, and early empires in turn gave rise to modern nation-states during the 19th century. The 20th century witnessed the rise of economic, political, and cultural globalization. Clearly, there has been a breakneck rush toward ever higher levels of cultural complexity over the past few centuries. Perhaps as a consequence of this pattern of historical evolution, cultural complexity has been used primarily to characterize premodern societies rather than to differentiate between modern nation-states.

Social scientists interested in the relationship between human culture and biology should keep this point clearly in focus: The modern complex human brain was in place far earlier than were the complex cultural and social manifestations that it permits. Evolutionary psychologists should be cognizant of this when inferring connections between current human social behavior and the biological apparatus of the mind.

2. CULTURAL COMPLEXITY DEFINED

Given that it is a central theoretical concept in the social sciences, especially anthropology, an explicit definition of cultural complexity is surprisingly hard to pin down. Carneiro provided one influential perspective, arguing that complexity focuses on society as a structural end product, whereas evolution focuses on the processes that gave rise to this. Following Spencer, Carneiro defined evolution as "a change from a state of relatively indefinite, incoherent homogeneity to a state of relatively definite, coherent heterogeneity, through continuous differentiations and integrations." Hence, cultural complexity would be characterized by coherent heterogeneity—a state of organized differentiation. Such a definition is necessarily vague in that cultural complexity is used primarily as a relative term to rank order various cultures and societies (and this can be done reliably) rather than standing alone in terms of essential meaning according to necessary traits or attributes.

Cultural complexity entails the presence of a multitude of social networks, social groups, hierarchies, power structures, roles, positions, and/or divisions. Whereas kinship is the primary organizational concept in simpler societies, complex societies have multiple levels of crosscutting ties, affiliations, responsibilities, membership, values, norms, and rights. Individuals pursue their own particular goals through social networks, and no two people have the exact same network or objectives. Perhaps the key difference between simple societies and complex societies is the presence of social power by which human actors manipulate others in pursuit of social goals. With social power comes the potential for social inequality, ranking, hierarchy, and social stratification. It is also possible, however, to have multiple divisions of horizontal rather than vertical power, that is, heterarchy rather than hierarchy. In such cases, there is a different but equal distribution of power.

3. THEORIES ON THE ORIGINS OF CULTURAL COMPLEXITY

A wide range of theoretical perspectives has been employed to understand how and why human cultures become complex. Progressivists such as Locke first opened the door to explaining differences in human societies without recourse to religious or mythological causes. By the mid-19th century, naturalists and social

scientists alike were profoundly influenced by Darwin's theory of natural selection. Social Darwinists such as Spencer and early anthropologists such as Tylor applied evolutionary concepts to the development of human society. Morgan and other unilineal cultural evolutionists believed simplistically that human societies naturally progressed through a series of stages usually tied to economic or technological patterns. Such thinkers, unfortunately, tended to invoke racial arguments to explain differences in cultures and their level of attainment. Nevertheless, such thinking had a strong influence on Marx in his theories of social evolution and economic materialism. Similarly, Freud was heavily influenced by unilineal evolution as he outlined the development of human societies from simple to complex through universal psychological patterns.

The early and mid-20th century saw two major perspectives on cultural complexity struggle for supremacy in both American and European social sciences. "Multilinear evolutionists" favored materialist explanations focused on economics, technology, energy, ecology, and the like. Unlike their 19th-century predecessors, however, these theories recognized numerous causal factors and possible cultural trajectories. Evolutionists rarely gave credence to ideology, symbolism, or the role of the individual. Cultural relativists, on the other hand, avoided generalizations and focused instead on the diffusion of cultural traits or functional descriptions of culture and social structure. The most prominent example in American cultural relativism is Boas. In Europe, social theory was led by schools of sociology (e.g., Durkheim, Mauss, Weber), structuralism (e.g., Levi-Strauss), and functionalism (e.g., Malinowski, Radcliffe-Brown, Evans-Pritchard). These perspectives often downplayed materialism and instead employed ideology, symbolism, and deep cognitive structures as causal factors. Some of these perspectives also attached great importance to the role of the individual as a cultural actor with agency.

The social sciences, however, did develop a synthesis view of cultural evolution and cultural complexity by the end of the 20th century. It recognizes that extreme views are unlikely to derive a parsimonious explanation for all cases. Mann has been a leading figure of the synthesis through his multidimensional social power theory, which stresses that individuals are agents that pursue their goals through overlapping networks of economic, military, political, and economic power. Social power is the means by which individuals acquire and manipulate the flow of information, energy, social status, wealth, and other resources.

Social scientists during the past few decades have had a more sophisticated view of cultural complexity, one that takes account of a variety of overlapping factors.

4. MEASURING CULTURAL COMPLEXITY

A recent review by Chick outlined three major approaches to the measurement of cultural complexity (a broader review of seven specific measures was developed earlier by Levinson and Malone). The three types of measures are highly correlated, and each has specific strengths and weaknesses. The approaches are (a) attribute counting procedures, (b) development scores based on the Standard Cross-Cultural Sample (SCCS), and (c) maximum settlement size.

The first measure is termed "culture by accretion." In this method, cultural complexity is assessed by an attribute checklist; the greater the number of attributes possessed, the greater the complexity score. As few as 50 and as many as 618 attributes have been used. The greater the number used, the more accurate the score and the greater the ability to differentiate between cultures at similar levels of complexity. If only a limited number is chosen (e.g., 50), items conforming optimally to Guttman scaling procedures are required.

Domains for the attributes (or traits) used by Carneiro in 1970 were economics, social organizations and stratification, political organization, law and judicial process, warfare, religion, art, tools–utensils–textiles, metalworking, watercraft and navigation, and special knowledge and practices. Sample items included the presence or absence of markets, military conscription, special religious practitioners, and temples. Most of the items were related to technological and organizational elements of culture rather than expressive ones, but the breadth of the attributes nonetheless allows researchers to empirically assess the level of cultural complexity.

Although comprehensive, reliable, and accurate, such measures have fallen out of fashion due to the difficulty in obtaining scores for such a large number of attributes and because the attributes chosen seem to be specific to cultures at certain stages of development. They measure increases in structure rather than growth in the size of a structure (e.g., a society with 7 judges is of equal complexity to a society with 7000 judges in this system). Because the most complex society studied by Carneiro was the Roman Empire, such domains as information technology or energy sources were not assessed in his scoring system.

The second approach, perhaps the best known among the approaches, is based on 10 indicators of development from the SCSS, with each being rated from 0 to 4 and largely focused on technology: writing and records, fixity of residence, agriculture, urbanization, technological specialization, land transport, money, population density, political integration, and social stratification. In 1973, Murdock and Provost reported a correlation of .95 between their scale and Carneiro's scale using a sample of 45 societies common to both studies.

Because this measure depends on rating the degree of sophistication of a few indicators rather than counting the presence or absence of many attributes, Chick called it "complexity by elaboration." The measure factors into two dimensions: social and technological complexity and the complexity of human ecology. Like Carneiro's measure, the sample of the SCCS has few complex modern societies, and so the top end of the scale might not differentiate well. Chick argued that the scale does not adequately assess expressive elements of culture such as arts and dance. But in 1972, Lomax and Berkowitz provided evidence that cultures synchronize activity in productive efforts and expressive performance.

Finally, and most succinctly, there is a measure of complexity by magnitude. Drawing from previous work showing a .97 correlation between Carneiro's scale and the logarithm of maximum settlement population, Naroll and Divale argued in 1976 that maximum settlement population is the single best measure of cultural complexity. The reason for this can be found in Carneiro's 1987 study of single-community societies (i.e., autonomous bands and villages). He reported that the organizational complexity of a village increases as the two-thirds power of its population. Villages that do not increase in organizational complexity as their population increases tend to fall apart. Extrapolating from this, it follows that in larger societies there should also be a mathematical relationship between maximum settlement size and the amount of organizational and technological complexity required to provide for and hold together these populations.

It would be difficult to apply this measure to non-sedentary populations, and it does not capture the types of cultural complexity afforded by globe-spanning mass communication and information technologies. It would also imply that Japan, with more than 30 million people living in agglomeration around metropolitan Tokyo, is the most complex society in the world today.

To summarize, despite difficulty in providing a precise conceptual definition of cultural complexity, and despite limitations in its operationalization, there is consistent evidence of the convergent validity for this concept. However cultural complexity is operationalized, its disparate measures are highly intercorrelated. This suggests that there is a common theoretical construct underlying these measures, however difficult it may be to articulate succinctly.

5. CULTURAL COMPLEXITY AND DIMENSIONS OF CULTURAL VARIATION

In relating cultural complexity to the dimensions of cultural variation typically used by psychologists, it is important to remember that the concept originates in anthropology. Most of the societies and cultures used by Carneiro or Murdock and Provost to establish their measures of cultural complexity either cannot or are unlikely to be represented in the corpus of data associated with psychology. This is not to say that the concept of cultural complexity is not useful to psychologists, but it does sound a warning that there is a severe restriction in range among the peoples studied by psychology, including cross-cultural psychology. Established measures of complexity are not especially sensitive to differences at the high end of complexity, which is the range drawn on by most psychologists.

On the other end, it is unclear how less complex societies would score on the dimensions of variation in subjective culture cherished by psychologists. Would hunters and gatherers tend toward individualism or collectivism? Hunters and gatherers certainly relied on kin groups for survival, but these tended to be loose coalitions that could change with the availability of resources or with social conditions. Although interdependence was a basic condition of survival, there were fewer social norms governing behavior and less enforcement of these norms than in more complex societies. In 1994, Triandis argued that complexity is related to affluence and to the existence of multiple subgroups with different norms (hence promoting individualism), but another view is that it is difficult to establish a clear linear relationship.

Lomax and Berkowitz, using both the Ethnographic Atlas and cantometric measures of song, provided evidence that the relationship is curvilinear, with group cohesiveness, group organization, and division

of labor being highest among cultures at middle levels of complexity. A similar argument could be applied to looseness–tightness. Data from an ecocultural approach found that only in the range from moderate (i.e., irrigated agriculture) to highly complex (i.e., industrial) societies is increased complexity associated with more looseness and less social conformity. In the range from hunter-gathers to the onset of agriculture, Berry found in 2001 that increased complexity, as operationalized by societal size, is actually associated with more social conformity (less looseness). At the high end, both the United States and Japan are incredibly complex but are currently used as instances of individualism, on the one hand, and collectivism, on the other.

There appears to be a more clear-cut relationship between cultural complexity and the second most frequently used dimension of cultural variation in psychology, that is, power distance. Put simply, cultural complexity affords power distance. By definition, the amount of differentiation of statuses, roles, affluence, and power with a society increases with its complexity. However, there is considerable latitude in how this differentiation is managed and perceived. For example, Americans aspire toward informality based on egalitarian ideals, whereas Japanese adopt more role-based formality in dealing with different statuses and roles. Although the real difference in incomes between people in upper and lower classes is greater in the United States, the psychological distance between the upper and lower parts of society is perceived to be lower among Americans than among Japanese. But over the broad range of human societies, given the tendency of individuals to exploit structural differences in power to their advantage, the perception of power distance should increase with increases in cultural complexity.

The measurement of cultural complexity is a reminder to psychologists that subjective culture is not the only viable definition of culture. The technological, material, organizational, and political–economic features of society taken by anthropologists to be indicators of cultural complexity may provide preconditions for the emergence of dimensions of variation in subjective culture found in cross-cultural psychology. From the other end, it may be extremely dangerous to assume that dimensions of cultural variation emerging from the analysis of samples from modern nation-states provide an adequate description of the full range of variability in subjective culture possible in human society.

6. GLOBALIZATION AND CULTURAL COMPLEXITY

The historical and archaeological record over the past 10,000 years can be read as a steady movement toward increasing levels of cultural complexity, with a massive acceleration during the past 200 years. Concurrent with increases in complexity has been a decrease in the number of independent polities in the world. There were more than 100,000 world polities circa 500 AD compared with the approximately 200 sovereign states represented in the United Nations today. With the onset of globalization, where information technology, high-speed travel, trade, and mass media have reduced geographic boundaries between societies, representing dimensions of cultural variation as dichotomies (e.g., individualism vs collectivism) is becoming increasingly problematic. It is now possible to speak of a global system and the cultural complexity of this system, with hybridity and contact zones between cultures being just as important as core values and practices within cultures.

In such a global system, it may be interesting to speculate about the ramifications of increased complexity. In the developing world, traditional cultures will continue to struggle with the onset of modernity. Collectivism and tightness in these cultures will give way before or hybridize with individualism and looseness. In the system as a whole, power distance will continue to increase, with immense concentrations of power and wealth being in the hands of global elites. In such an emerging system, words such as "individualism" and "looseness" might be inadequate to describe the values and roles of people at the top end of human society. The influence of economic elites—whatever their nationality—on the evolution of global cultural complexity will be immense and largely unprecedented in human history.

Of course, social scientists are aware that cultural evolution is not a one-way street. Cultural devolution, or the reduction of complexity, has also been documented and is usually caused by political or ecological catastrophes. This is not an unforeseen possibility for observers of the global system. Whereas the conceptualization and measurement of cultural complexity has been focused on simpler societies, future research might undertake to describe cultural evolution at the high end of complexity as social scientists seek to describe the interaction between a global supersystem and its constituent parts.

See Also the Following Articles

Acculturation ■ Conformity across Cultures

Further Reading

Berry, J. W. (2001). Contextual studies of cognitive adaptation. In J. M. Collis, & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement* (pp. 319–333). Mahwah, NJ: Lawrence Erlbaum.

Carneiro, R. L. (1970). Scale analysis, evolutionary sequences, and the rating of cultures. In R. Carroll, & R. Cohen (Eds.), *A handbook of method in cultural anthropology* (pp. 834–871). Garden City, NY: Natural History Press.

Carneiro, R. L. (1987). The evolution of complexity in human societies and its mathematical expression. *International Journal of Comparative Sociology*, *28*, 111–128.

Chick, G. (1997). Cultural complexity: The concept and its measurement. *Cross-Cultural Research*, *31*, 275–307.

Harris, M. (1968). *The rise of anthropological theory: A history of theories of culture*. New York: Thomas Y. Crowell.

Hermans, H. J. M., & Kempen, H. J. G. (1998). Moving cultures. *American Psychologist*, *53*, 1111–1120.

Levinson, D., & Malone, M. J. (1980). *Toward explaining human culture: A critical review of the findings of worldwide cross-cultural research*. New Haven, CT: Human Relations Area Files Press.

Liu, J. H., & Allen, M. W. (1999). Evolution of political complexity in Maori Hawke's Bay: Archaeological history and its challenge to intergroup theory in psychology. *Group Dynamics*, *3*, 64–80.

Lomax, A., & Berkowitz, N. (1972). The evolutionary taxonomy of culture. *Science*, *177*, 228–239.

Mann, M. (1986). *The sources of social power*. Cambridge, UK: Cambridge University Press.

Murdock, G. P., & Provost, C. (1973). Measurement of cultural complexity. *Ethnology*, *12*, 379–392.

Murdock, G. P., & White, D. R. (1969). Standard cross-cultural sample. *Ethnology*, *8*, 329–369.

Naroll, R., & Divale, W. T. (1976). Natural selection in cultural evolution: Warfare versus peaceful diffusion. *American Ethnologist*, *4*, 97–128.

Triandis, H. (1994). *Culture and social behavior*. New York: McGraw–Hill.

Cultural Psychology

Patricia M. Greenfield
University of California, Los Angeles, California, USA

Heidi Keller
University of Osnabrueck, Osnabrueck, Germany

GLOSSARY

apprenticeship The interactive process by which cultural experts help novices acquire knowledge in a particular domain.

coconstruction Creation of shared meanings or shared activities through an interactive communication process.

community of practice A group that regularly comes together to participate in shared activities.

cross-cultural psychology The study of similarities and differences across cultures in human behavior, cognition, and emotion, with culture usually being an independent variable. It differentiates etic and emic approaches. Etic approaches imply the transfer of concepts and methodologies from one culture to other cultures; emic approaches imply the development of culturally sensitive concepts and methodologies that are derived from within each culture to be studied. Cross-cultural psychology was first linked to cognitive psychology and cognitive development and later to social psychology and social development.

ethnography An anthropological method of field research through participant observation. The researcher utilizes a participatory role as a position from which he or she can make observations in a particular cultural context. Ethnography is generally carried out by researchers who do not belong to the ethnic community under study but who spend a considerable amount of time within the ethnic community and usually speak the local language. Ethnography covers observational methods, interviews with cultural informants, and focus group discussions.

ethnotheory A folk or lay theory of some part of the everyday world.

expert An expert has acquired advanced cultural knowledge in a particular domain.

independence The self-construal prevailing in large-scale, urban industrialized cultural communities, such as the United States or Western Europe. It denotes a cultural deep structure of a self-contained and bounded self that is separate from and competes with social others. The deep structure becomes instantiated in locally appropriate phenotypical expressions.

indigenous psychology The study of local conceptions of humans, their development, and their psychology. Indigenous approaches express dissatisfaction with the ethnocentric Euro-American view of psychology, with its claims of universality. There are two conceptual approaches, one aimed at describing variability within a universal framework (the unity of mankind) and the other aimed at establishing diverse psychologies.

interdependence The self-construal prevailing in small-scale, rural, subsistence cultural communities in Asia, Africa, and Latin America. It denotes a cultural deep structure of an interrelated self that maintains harmonious and hierarchically structured relationships with social others. The deep structure becomes instantiated in locally appropriate phenotypical expressions. Historical transitions and multicultural societies may support expressions of both

independence and interdependence in the definition of the self.

macrodevelopment Directed change over a relatively long timescale, such as years.

microdevelopment Directed change over a relatively short timescale, such as hours or days.

novice A novice is at the beginning stages of acquiring cultural knowledge in a particular domain.

parental ethnotheories The social cognitions that humans hold with respect to the concept of a child, the concept of a parent, and the developmental process. They are organized as belief systems that pertain to different domains. Parental ethnotheories are shared among members of cultural communities. They have intuitive, nonconscious components and serve as frameworks for parental behaviors and actions.

Social interaction among beings biologically primed for culture creates culture for the group and for the individuals in it. Culture is viewed as a socially interactive process with two main component processes: the creation of shared activity (cultural practices) and the creation of shared meaning (cultural interpretation). Shared cultural knowledge and modes of cognition are the result of processes of coconstruction. The creation, acquisition, transmission, and use of culture are psychological and interactional processes. It is the study of these processes that is the paradigmatic subject of cultural psychology.

1. INTRODUCTION

Shared activities are adapted to survival in a particular environment; they are the material side of culture. A key aspect of shared activity is that it involves goal-directed action; this point is particularly emphasized in the German tradition of cultural psychology. Shared meaning is the symbolic side of culture. A key aspect of shared meaning is interpretation, the notion that culture provides a lens through which we see and interpret (evaluate) the world. Empirical study in cultural psychology focuses on these two components of cultural process—shared meaning and shared activity.

Both components are cumulative in nature because culture is created by processes that occur between, as well as within, generations. Meanings and activities not only accumulate but also transform over time. This cumulative and temporal characteristic of culture creates the necessity for historical methods of studying culture as psychological process. It also creates the

necessity for developmental methods for studying the transmission and transformation of culture from the older to the younger generation. Equally important (albeit understudied) are developmental transitions from being a novice socialized by experts to being an expert socializing novices. Cultural psychology does not conceptualize the adult as a static entity with set cultural characteristics but rather attempts to study the diachronic processes of socialization and development that constitute cultural learning and apprenticeship.

As a species, humans are biologically primed to create, acquire, and transmit culture; culture is the primary mode of human adaptation. Because of the centrality of adaptation as a concept, cultural psychology aims to study meanings and activities in context rather than in the more decontextualized setting of the laboratory experiment. Cultural psychology is an interdisciplinary specialty that integrates psychology and anthropology.

2. THE CONCEPTION OF CULTURE IN CULTURAL PSYCHOLOGY: CONCEPTUAL AND METHODOLOGICAL ISSUES

Cultural psychology can be understood by contrast with two other approaches to the relationship between psychology and culture—cross-cultural psychology and indigenous psychology. Although all three are fuzzy concepts with partially overlapping exemplars, the paradigmatic instances of the three approaches are quite distinctive; these paradigms are central to the descriptions that follow.

2.1. Cultural Psychology Compared with Cross-Cultural Psychology

The goal of cross-cultural psychology is to separate the universal from the culture specific in human behavior. Cross-cultural psychology has a universalistic bias: Culture can be decomposed, its different layers can be peeled off, until the existential and universal human nature remains. In cross-cultural psychology, culture is generally operationalized as an antecedent or independent variable to behavior. Often, the independent variable is the label of a particular national group (e.g., a comparison of Chinese and U.S. participants). Occasionally, as in Harry Triandis's 1980 definition, culture is also conceived as a dependent variable, with

behavior or experience as the independent variables. Whether an independent or a dependent variable, culture is implicitly viewed as being outside of and apart from the individual.

In cultural psychology (original terminology from Price-Williams), culture is, in contrast, seen as inside the individual. Culture is "a way of knowing, of construing the world and others" (Bruner, 1993, p. 516). Through processes of interaction and communication, these construals acquire a certain degree of intersubjectivity or shared meaning. Shared knowledge and shared meanings generate a set of everyday practices that also define culture in terms of shared activities. Thus, as Jahoda points out, culture and behavior, culture and mind are viewed as indistinguishable.

Accordingly, cultural psychology derives its problems and procedures from an analysis of the nature of culture, with an important focus on everyday life. Whereas the methodological ideal of the cross-cultural psychologist is to carry a procedure established in one culture, with known psychometric properties, to one or more other cultures, to make a cross-cultural comparison, the methodological ideal of the cultural psychologist is for problems and procedures to flow from the nature of culture, both in general and in specific. For example, in their 1981 book, Scribner and Cole focused on literacy as an important cultural element, a way of interpreting and representing the world (the nature of culture in general). They proceeded to find out what each of three different literacies meant to the Vai people of Liberia (the nature of culture in specific) and to study the cognitive consequences of these literacy practices.

The cultural psychology of Scribner and Cole traces its intellectual roots to the work of Russian psychologists Vygotsky and Luria. In the early 1930s, soon after the Russian revolution, Vygotsky and Luria studied the logical processes of collectivized and noncollectivized central Asian peasants. Their stimuli were two kinds of syllogism: "The contents of some were taken from the concrete, practical experience of the villagers; the contents of others bore no relation to familiar, practical life" (Cole & Scribner, 1974, p. 161). The point of their study was the cognitive difference in the way the two groups of peasants handled the two kinds of syllogism. Again, syllogisms are an aspect of human logic (culture in general), whereas the everyday life of the peasants (culture in specific) was a key source of problems to solve. These examples from Scribner and Cole and Vygotsky and Luria not only illustrate the methods of cultural

psychology but also illustrate the importance of cognition in the origins of cultural psychology.

Thus, cultural psychologists study cultural processes directly—for example, literacy or syllogisms as cultural forms and products. In comparison with cross-cultural psychologists, their research designs rely much less on what the anthropologist Beatrice Whiting called "packaged" variables. A packaged variable is an index of culture as an antecedent or independent variable. An example of a packaged variable is the labels "Japanese" or "German." In principle, each of these labels "packages" a whole suite of shared cultural characteristics and processes. When cultural psychologists make comparisons, they do not use packaged variables. Instead, they use much narrower groupings, such as collectivized vs uncollectivized peasants in Luria and Vygotsky's work.

Some research falls on the border between cultural and cross-cultural psychology. In the 1960s, for example, Greenfield brought Piagetian cognitive tasks to the Wolof of Senegal, an approach belonging to cross-cultural psychology. However, she did not compare Senegalese to North Americans, which would have been a cross-cultural approach. Instead, she looked at developmental trajectories within each culture, as well as compared children from different ecologies within the Wolof world: urban vs rural and schooled vs unschooled. She also examined the Whorfian hypothesis by doing a semantic analysis of various domains in Wolof and French. This was another way in which research procedures flowed from the nature of the culture—in this case, the linguistic aspects of culture.

The cultural psychological perspective implies that research is conceptualized as a communication process. Communication with the people of the study community in their own language is a prerequisite for cultural analyses of shared activities and shared meanings. The anthropological notion of ethnography as a methodological concept is a necessary first step in order to bring the importance of lived experience in a cultural place from ground to figure. In order to study cultural processes from the participants' own points of view, cultural psychology has, as Greenfield pointed out in 1997, developed its own tool kit of methods. Some, like ethnography and discourse analysis, have been adapted from anthropology. With other methods, such as the use of video, cultural psychology has followed the leadership of developmental psychology as a discipline. (Later, exemplars of different research programs are introduced, noting different cultural methodologies for different research problems.)

2.2. Cultural Psychology Compared with Indigenous Psychology

Although indigenous psychology and cultural psychology clearly have independent origins, they share the notion that the prime subject of study is the subject's creation of meaning systems, particularly systems that are shared or normative within a defined cultural group. In different ways, both traditions have recognized that psychological theories are important aspects of shared cultural meaning. Millers, writing from the vantage point of a cultural psychologist, has, on a theoretical level, asserted the cultural grounding of all psychological theory. This cultural grounding of theory has been a strong motive, if not an explicit metatheoretical statement, for indigenous psychology.

The unique contribution of indigenous psychology is the notion that psychological concepts and psychological theory, not just data-collection techniques, should be developed within each culture. Unlike indigenous psychology, the empirical research tradition of cultural psychology has not been based on formal psychological theories with culture-specific origins. Nevertheless, cultural psychology has increasingly in recent years made ethnotheories (i.e., folk theories) of psychological functioning and development a subject for empirical investigation—for example, in the collaborations of Shweder and Bourne and Harkness and Super.

Indigenous psychology, however, aims to go one step further: The goal of indigenous psychology is to take informal folk theories of psychological functioning and formalize them into psychological theories. Cultural psychology arrived at the empirical study of folk theories, including folk theories of psychological development. Indigenous psychology has, in turn, moved ethnopsychology from an object of empirical study to a source of formal psychological models. In other words, indigenous psychologists have taken steps to translate ethnotheories of psychology into formal theories of psychology and, from these theories, to conduct empirical psychological studies. This goal is incredibly important for psychology as a whole, on both a metatheoretical and a metamethodological level.

However, cultural and indigenous psychology also differ in some respects. Unlike indigenous psychology, which, as Sinha points out, was born as an attempt to decolonize the mind—in fact, many of its pioneers and proponents inhabit former British colonies—cultural psychology still constitutes a "crossing over" into someone else's culture by the investigator. Also, unlike indigenous psychology, its connection to cross-cultural psychology has tended to be erratic and ambivalent.

On the other hand, because of its own origins with an emerging scientific elite in developing nations, indigenous psychology, especially in East Asia, tends, on the whole, to privilege elite populations (university students) as subjects of study and culture change as a research topic. Cultural psychology, in contrast, tends to give a great deal of attention to relatively stable subsistence village cultures; for example, well-known studies have been done on subsistence groups in Liberia, Morocco, Guatemala, India, and Mexico.

2.3. Comparing Indigenous, Cross-Cultural Psychology, and Cultural Psychology

Although indigenous psychology shares the spirit of cultural psychology, its methods (in practice, although not in principle) tend to resemble those of cross-cultural psychology. That is, indigenous psychology most often utilizes standard psychological methodology such as questionnaire formats and tends to study variables rather than processes. For example, Yang and Bond used rating scales and adjective lists in their 1990 study of indigenous Chinese personality constructs.

There are some notable exceptions. For example, Enriquez, the founder of indigenous psychology in the Philippines, has utilized an indigenous (and group-oriented) method of social interaction as a means for data collection. However, by and large, indigenous psychology, unlike cultural psychology, has not experienced the influence of method and theory from anthropology and developmental psychology.

Why has cultural psychology developed in the way that it has? To answer this question, we turn to a study of its sociopolitical roots.

3. SOCIOPOLITICAL ROOTS OF CULTURAL PSYCHOLOGY

Cultural psychology has a long tradition and a rather young history in psychology. In the context of defining the place of psychology in the systematics of sciences, in 1920 Erich Stern presented a view that could not be more modern. For him, psychology is inevitably a cultural science since any attempt to understand any psychological phenomena needs to take into account the social cultural environment, ontogenetic history,

and ancestral heritage. Culture represents the legacy of preceding generations as expressed in the dispositions, the consciousness, and the psychology of each living individual whose plasticity allows change in order to adapt to changing surroundings. His thoughtful and comprehensive treatise also disqualified experiments as a tool to understand psychological processes.

The modern conception of cultural psychology developed without connection to such ancestors. It grew out of liberal policy in a world increasingly dominated by the Cold War after World War II. The notion was to help the Third World develop in order to make sure that, on the political level, Third World countries did not fall to Communism. Thus, John Gay and Michael Cole, funded by the Ford Foundation, began their cultural psychological research in Liberia as part of the African Education Program of Educational Services, Inc. of Watertown, Massachusetts. Daniel Wagner originally went to Morocco with the U.S. Peace Corps. Patricia Greenfield's dissertation research in Senegal in 1963 was funded by a grant (awarded to Jerome Bruner at Harvard University) from the U.S.-based Ford Foundation to develop the Institute of Pedagogical Studies at the University of Dakar in Senegal.

Initially independent from this U.S. branch of cultural psychology, a German school of cultural psychology emerged in Saarbruecken. It was founded by Ernest Boesch and further developed by Lutz Eckensberger. These scholars also went to so-called Third World cultures on the basis of politically motivated developmental aid programs. Ernest Boesch went to Thailand using grants from the UNESCO in order to set up an institute; Lutz Eckensberger went to Afghanistan using grants from the former German Ministry of Developmental Aid in order to train teachers in vocational schools. The initial contact with other and very different cultures thus became a major factor in conceptualizing cultural psychologies.

On the institutional level, the Department of Social Relations at Harvard, with its integration of psychological, sociological, and anthropological methods and theories, provided the educational foundation of modern cultural psychology and was particularly crucial in the merger of psychology and anthropology. The actual term cultural psychology seems to date from 1969, when two anthropologists, DeVos and Hippler, wrote an article titled "Cultural Psychology: Comparative Studies of Human Behavior" for the second edition of Lindzey and Aronson's *Handbook of Social Psychology*. A second use of the term cultural psychology emerged in 1980 when Douglass Price-Williams, a psychologist

serving as a member of an anthropology department, wrote an article titled "Toward the Idea of Cultural Psychology: A Superordinate Theme for Study." Hence, cultural psychology has from the beginning represented the meeting of psychology and anthropology.

On the intellectual level, cultural psychology has grown out of dissatisfaction with the universalism and decontextualized methodology of psychology in general and cross-cultural psychology in particular. On the other hand, it has also grown out of anthropology's wish to deal with the person, not merely the culture as a supraindividual envelope. These motivations, from Cole, Shweder, and Bourne in the United States, contrast with that of German cultural psychology. The latter has, in reaction to the Nazi Terror, been driven by a strong desire for a humanistic psychology or even a more human world. These qualities seemed lacking in behavioristic approaches that became mainstream during that time.

4. EXEMPLARS OF RESEARCH PROGRAMS

4.1. Parental Ethnotheories: An Example of Shared Meaning

The analysis of parental ethnotheories about development has become a vital area of research. Parental ethnotheories are shared beliefs about the goals of child development and the socialization practices that will achieve these goals. According to Harkness and Super, they specify how to become a competent adult in a particular environment. Parents throughout the world hold specific beliefs about proper care and handling of small children. These parental ethnotheories express conceptions of the nature of children, parenting, and development. In this way, ideas about child care practices are related to developmental goals, as defined by the cultural environment. Parental ethnotheories thus link socialization practices to cultural values. They also link the socialization agendas of one generation to the next. Parental ethnotheories generate child-rearing practices that, in turn, create the cultural adult. This cultural adult, in turn, participates in the socialization processes of the next generation.

Although interindividual differences are prevalent with respect to parenting goals and practices in every culture, members of cultural communities can be viewed as groups of individuals who coconstruct a shared reality in different domains of life. According

to Markus and Kitayama, two basic contrasting realities are the value systems that define independent and interdependent construals of the self. These, in turn, as Greenfield and Cocking have shown, generate corresponding developmental goals. These goals are actualized through different socialization practices and parenting styles during different stages of development. Parental ethnotheories can thus be conceived of as the mediating links between these cultural metamodels and behavioral contexts and practices.

Keller and collaborators analyzed parental ethnotheories in cultural communities found to differ in their orientations toward independent and interdependent construals of the selves: West African Nso farmers (interdependent orientation) and middle-class northern Germans (independent orientation). Parental ethnotheories were assessed as reactions to videotaped prototypical interactional situations from these different cultural communities. (This procedure has also proved revealing of cultural patterns in research on teacher and classroom evaluation.) Particularly because these reactions were assessed in a group setting, in which consensus was generally negotiated, it can be assumed that the respondents use their shared cultural framework with all the cultural standard operating procedures and unstated cultural assumptions. Because of the use of actual interactions as stimuli, this procedure further allows one to address the issue of the belief–behavior dissonance. Comments concerning deviations from normative practices reveal the valued practices and beliefs.

These analyses revealed results that are in line with the expectations of different cultural orientations with respect to independence and interdependence. The Nso women comment on parental behaviors that can be related to interdependent developmental goals such as primary care and body contact and subscribe to a physical stimulation model of infancy, in which their special technique of motor stimulation is supposed to foster growth and development. The German women, through their cultural lenses, value face-to-face communication, exclusive attention toward the baby, object play, and language. They subscribe to a communication model of infancy, in which communication occurs between two bodies that are physically separate and the flow of the interactional exchange is supposed to be shared by (quasi) equal partners, thus supporting independent socialization goals.

The evaluations of the German and the Nso women are independent of the origin of the videotapes. That is, the Nso cultural lens interprets physical stimulation practices positively when participants see them on videos from either culture. In contrast, the Nso cultural lens provides a negative evaluation when these stimulation practices are missing in the videos of either culture. Similarly, German mothers positively evaluate instances of the communication model and negatively criticize their absence on both the Nso and German tapes. However, physical stimulation behaviors are more prevalent in the Nso videos, whereas communication behaviors are more prevalent in the German videos. Thus, the cultural values manifest in the ethnotheories are reflected in socialization practices favored by each cultural group.

Ethnotheories link to socialization practices, which then link to children's behavioral development (e.g., the timing of mastering developmental tasks). Having experienced distinct and frequent body stimulation, the Nso babies develop the gross motor milestones of sitting, standing, and walking significantly earlier than the German babies. On the other hand, the experience of a rich language environment, face-to-face interactions with nonverbal and verbal contingencies, and object play instigate an earlier onset of language in German babies compared to Nso babies. The fact that the Nso community does not value the early onset of language as much as the German cultural community is expressed in a statement of a Nso woman, commenting on the language use of a German mother interacting with her baby: "For them, every child needs to be intelligent." The cultural timing of developmental milestones organizes subsequent developmental pathways in coherent and systematic ways so as to achieve the developmental goals valued in each cultural environment. The mothers who engage in these child-rearing practices were themselves the object of the corresponding socialization practices when they were children. It is believed that childhood socialization is perhaps the most conservative or persistent part of any culture. On the other hand, with changing sociocultural and especially socioeconomic environments, cultural practices also change. However, research shows that practices change long before attitudes and beliefs.

4.2. Apprenticeship: The Historical Transformation of Shared Practice

As Barbara Rogoff and Jean Lave note, apprenticeship is the participatory structure by which more expert members of a culture induct less expert members into the normative activities and practices of a culture. Corresponding to the two basic types of ethnotheory,

one idealizing the independent individual and one idealizing the interdependent individual, are two basic styles of apprenticeship. The independent style of apprenticeship utilizes more trial-and-error strategies, learner initiative, and division of labor. The interdependent style utilizes more teacher guidance, observational learning, and interdependent problem solving. The former is well adapted to a traditional agrarian setting in which information and resources flow from the older to the younger generation with relatively little change. The latter is well adapted to a commercial economy in which innovation and novelty are valued and the younger generation is more independent of the older.

However, the styles of apprenticeship are not constant; they are socially constructed in response to environmental conditions. Therefore, changes in economic conditions should induce a shift in mode of cultural transmission. This notion was tested by Greenfield, Maynard, and Childs in a video study of weaving apprenticeship in a Maya community at two points in time; the two waves of data collection were separated by two decades. At the second point of data collection, the weaving learners from the first wave had grown up and become mothers whose daughters were now learning to weave.

In both cohorts, researchers microanalyzed weaving videos to understand the interaction between learner and teacher as apprenticeship progressed. The first generation of weaving teachers, mainly mothers, utilized the interdependent style of apprenticeship as their daughters learned to weave. Two decades later, after a period in which the economy moved away from subsistence and toward commerce, the mode of apprenticeship had indeed shifted. Microanalysis of the videos revealed the appearance of a new, more independent style of apprenticeship. Weaving learners were displaying more initiative in the apprenticeship process, peer teachers (compared with teachers from the older generation) had become more prevalent, and there was less cooperative weaving between learner and teacher. This shift was concentrated, moreover, in just those girls who, with their mothers, were most involved in textile commerce.

This historical transition in the apprenticeship process was reflected in changes in resulting weaving practices. What girls and women wove had also changed. The concept of community creativity, in which a few fairly uniform designs identified members of a cultural community, was joined by a concept of individual creativity. Design innovation and individuation were now apparent throughout the woven textiles. Shared practices had changed in the direction of a new cultural model.

The dimension of shift was that dimension most basic in the cultural differentiation of ethnotheories of development discussed previously. This parallel suggests a close conceptual relationship between the analysis of cultural meanings and the analysis of cultural practices.

4.3. Connecting Shared Practice to Shared Meaning

Although development and the analysis of developmental processes are an integral part of cultural psychology, not all research programs address ontogenetic, intergenerational, or historic development directly. Adult cultural experience is also a focus of cultural psychology. An important theoretical issue is the extent to which shared practice results in shared cultural meaning. In the German tradition of cultural psychology, Lutz Eckensberger and colleagues used different microcommunities of practice to explore this issue with adults. The researchers interviewed different samples of people who had different views on a coal plant that was supposed to be built. The groups included people living in the area, experts on coal plants, environmental activists, and politicians. Each group could be considered to share activities and practices in relation to coal plants in general and this plant in particular.

The investigators used interviews to assess the moral meanings constructed by the different groups concerning the coal plant. Their interest was in the structure (e.g., complexity) of the reasoning rather than in its content. Findings revealed that each group had a typical pattern of moral reasoning about the coal plant. Thus, there was a link between activities common to each group in relationship to the plant and level of moral reasoning. The pattern of results indicates a connection between shared activities and the construction of shared structures of moral meaning.

5. APPLICATION OF CULTURAL PSYCHOLOGY

As these examples show, cultural psychology can be applied to analyzing the everyday realities of members of different cultures and subcultures. However, there is another level of application that is also possible: applying findings to social interventions. We present one example. An implication of the modern-day multicultural reality in countries such as the United States or Germany is seen at

school. Multicultural immigration opens up the possibility of different cultural values among students, between students and teachers, and between home and school. An applied project, "Bridging Cultures," began with basic research by Greenfield, Quiroz, and Raeff documenting cross-cultural value conflict between Latino immigrant families and schools. Immigrant parents were generally much more collectivistic in their orientation to child socialization than were their children's teachers. In collaboration with a group of teachers, Greenfield, Trumbull, Rothstein-Fisch, and Quiroz utilized this research to help teachers and schools understand home culture and school culture in order to create educational "bridges" between them. The result of this intervention was greater home–school harmony in educational practices and classroom management. Shared practices and meanings were created between home and school.

6. CONCLUSIONS: AN EXAMPLE OF CULTURAL PROCESS

Both shared activities and shared meanings are intrinsic to the human mode of adaptation for survival. They are two facets of shared cultural knowledge and the primary subject matter of cultural psychology. An example will elucidate how shared activities and shared meanings are coconstructed in communities of practice; such coconstruction is the paradigmatic cultural process. The example illustrates that both material culture and symbolic culture result from processes of cultural coconstruction.

Following the Los Angeles earthquake of 1994, many of the material supports of everyday life, such as water, electricity, and roads, were destroyed. In small groups and through the media, people developed new shared knowledge concerning survival activities, such as where to get water, how to circumvent damaged roads to get from point A to point B, and methods for detecting leaking gas. Expertise was shared with novices, such as when a contractor showed his neighbors how to turn off their gas or a ham radio operator provided news of the location and magnitude of the earthquake in the absence of electricity. The nature of culture as a tool for organizing everyday life, a notion developed by anthropologist Thomas Weisner, was quite apparent.

Symbolic communication, through both language and visual media, is a critical means by which social sharing takes place; communication processes were quite intense during this period of adapting to the physical conditions created by the earthquake. As a result, new shared activities that enhanced physical survival were created through cultural processes of social interaction.

Simultaneously, shared meanings were also created to rationalize and understand the events that had taken place. Like shared activities, shared meanings arose through communication processes. One shared meaning that developed was the custom of asking people how they fared in the earthquake; the normative reply was, "I was fortunate." The search to create shared meaning for a stunning physical event was particularly apparent when, a few days after the earthquake, a local public affairs radio show host convened clergy from many religions to ask them about the larger meaning of the earthquake. His question was, "Did God send the earthquake to punish Los Angeles?" Clearly, adaptation to the aftermath of the earthquake could not be reduced to a process of adapting to physical conditions; the interpretation of these conditions—that is, processes of creating meaning—were part and parcel for the shared culture that developed in response to the earthquake.

This example provides a model of processes that are assumed to occur whenever a new member of society is born: the creation of shared knowledge, activities, conventions, and meanings through communication and social interaction. This microdevelopmental example of culture re-creation by adults occurs in each generation in the macrodevelopmental processes of children. The example is also a model and metaphor of culture change provoked by new ecological conditions. Finally, this example illustrates the potential for cultural variability as a response to different ecological conditions. This is the stuff of cultural psychology.

See Also the Following Articles

Child Development and Culture ■ Community Psychology ■ Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Cultural Complexity ■ Diverse Cultures, Dealing with Children and Families from ■ Family and Culture ■ Indigenous Psychologies ■ Intergroup Relations and Culture ■ Motivation and Culture ■ Perception and Culture ■ System Safety ■ Values and Culture

Further Reading

Berry, J. W., Poortinga, Y. H., Pandey, J., Dasen, P. R., Saraswathi, T. S., Segall, M. H., & Kagitcibasi, C. (1997). *Handbook of cross-cultural psychology* (2nd ed.). Boston: Allyn & Bacon.

Boesch, E. E. (2001). Symbolic action theory in cultural psychology. *Culture and Psychology.* 7, 479–483.

Bruner, J. (1993). Do we "acquire" culture or vice versa? *Behavioral and Brain Sciences*, *16*, 515–516.

Cole, M. (1995). Culture and cognitive development: From cross-cultural research to creating systems of cultural mediation. *Culture & Psychology*, *1*, 25–54.

Cole, M. (1996). *Cultural psychology: A once and future discipline.* Cambridge, MA: Harvard University Press.

Eckensberger, L. H., Döring, T., & Breit, H. (2001). Moral dimensions in risk evaluation. *Research in Social Problems and Public Policy*, *9* (Special issue).

Enriquez, V. G. (1993). Developing a Filipino psychology. In U. Kim, & J. Berry (Eds.), *Indigenous psychologies: Research and experience in cultural context* (pp. 152–169). Newbury Park, CA: Sage.

Greenfield, P. M. (1997). Culture as process: Empirical methods for cultural psychology. In J. W. Berry, Y. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology. Vol. 1: Theory and method* (pp. 301–346). Boston: Allyn & Bacon.

Greenfield, P. M., & Cocking, R. R. (1994). *Cross-cultural roots of minority child development.* Hillsdale, NJ: Erlbaum.

Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology*, *54*, 461–490.

Greenfield, P. M., Maynard, A. E., & Childs, C. P. (2003). Historical change, cultural learning, and cognitive representation in Zinacantec Maya children. *Cognitive Development*, *18*, 455–487.

Harkness, S., & Super, C. M. (1996). *Parents' cultural belief systems: Their origins, expression, and consequences.* London: Guilford.

Keller, K., Völker, S., & Yovsi, R. D. (In press). Conceptions of good parenting in Cameroonian Nso and Northern Germans. *Social Development.*

Kim, U., & Berry, W. (Eds.). (1993). *Indigenous psychologies: Research and experience in cultural context.* Newbury Park, CA: Sage.

Miller, J. G. (2003). The cultural grounding of social psychological theory. In A. Tesser, & N. Schwartz (Eds.), *Blackwell handbook of social psychology. Vol.1: Intrapersonal processes.* Oxford, UK: Blackwell.

Romero, M. E. (1994). Identifying giftedness among Keresan Pueblo Indians: The Keres study. *Journal of American Indian Education*, *34*, 35–58.

Scribner, S., & Cole, M. (1981). *The psychology of literacy.* Cambridge, MA: Harvard University Press.

Sinha, D. (1997). Indigenizing psychology. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology* (2nd ed., pp. 129–170). Boston: Allyn & Bacon.

Trumbull, E., Rothstein-Fisch, C., Greenfield, P. M., & Quiroz, B. (2001). *Bridging cultures between home and school: A guide for teachers.* Mahwah, NJ: Erlbaum.

Cultural Syndromes

Harry C. Triandis

University of Illinois at Urbana–Champaign, Champaign, Illinois, USA

GLOSSARY

collectivism A cultural pattern where the importance of the in-group is paramount; people think of themselves as aspects of groups, give priority to in-group goals, behave according to the norms of the in-group, and do not leave the in-group even if they are dissatisfied with it.

cultural complexity The contrast between simple cultures (e.g., hunters and gatherers) and information societies.

cultural tightness Cultures that have a very large number of rules and norms about social behavior and in which people are severely punished when they do not follow these rules and norms.

cultures high in diffusion Cultures that tend to think holistically, not differentiating among various aspects of situations.

cultures high in specificity Cultures in which people think analytically, distinguishing among various aspects of situations.

egalitarian cultures Cultures that emphasize equality.

hierarchical cultures Cultures that are organized according to status levels, where the dictates of those at the top of the social structure are given much weight and the opinions of those at the bottom of the social structure are not given much importance.

individualism A cultural pattern where the importance of the individual is paramount; people think of themselves as independent of groups, give priority to their personal goals, behave according to their attitudes, and easily leave groups with which they are dissatisfied.

masculine cultures Cultures that emphasize the differences between men and women and in which men rarely do the jobs that women do.

outcome cultures Cultures that emphasize the outcomes of action.

particularism Cultures in which behavior depends on the attributes of the person with whom one is interacting.

process cultures Cultures that emphasize how things are done.

universalism Cultures in which behavior is based on general principles rather than on the specific attributes of the person with whom one is interacting.

A cultural syndrome is a pattern of shared attitudes, beliefs, categorizations, self-definitions, standard operating procedures, unstated assumptions, norms, roles, and values that is organized around a theme. The theme is shared among those who speak a particular language, during a specific historical period, and in a definable geographical region. Cultural syndromes provide information about cultural differences. The number of existing cultural syndromes is unknown. However, some examples of cultural syndromes that have been identified are provided in this article.

1. EXAMPLES OF CULTURAL SYNDROMES

The following subsections describe five cultural syndromes that have been studied better than other cultural syndromes.

1.1. Cultural Complexity

The contrast between hunters and gatherers and information societies is vast. Simple cultures have few members, whereas complex cultures have many members. The number of people who constitute a cultural group is only one clue. Hunters and gatherers usually consist of bands of approximately 50 individuals who are related to each other through marriage or some other ceremonies. Information societies consist of millions of individuals who are sometimes related to each other through the Internet. In between, there is a very large number of levels of complexity such as "slash and burn" agriculture, regular agriculture, and industrial cultures.

Complexity is associated with the presence of writing systems, the presence of records, fixity of residence, agriculture, urban settlements, technical specializations, modes of transportation other than walking, money, high population density, many levels of political integration, many levels of social stratification, and the presence of different ways in which to make a living. For example, the number of occupations in simple cultures is small, whereas approximately 250,000 occupations have been identified in complex cultures. The number of choices available to individuals in simple societies is small, whereas it is vast (e.g., dozens of kinds of mustard) in complex societies. The number of standards for judging religious, economic, political, educational, social, and aesthetic phenomena is small in simple societies, whereas it is very large in complex societies. Thus, the organizing theme of this cultural syndrome is complexity.

1.2. Tightness

In tight cultures, there are many norms, standards, and rules for behavior. Those who do not do what is specified by these norms are criticized, punished, or even killed. In loose cultures, there are few norms, and people tolerate deviations from these norms. When a person does not do what is expected, people simply smile and say that it does not matter. Again, in between these two extremes, there are thousands of levels of tightness. The organizing theme is doing what is specified by in-group norms.

Conformity is high in tight cultures. Extreme examples of tight cultures are the Taliban in Afghanistan and North Korea. Tight cultures emerge in societies that are relatively homogenous, so that people can agree concerning what norms are to be used as guides for behavior. They also emerge when there is much interdependence required to make a living. For instance, to get a crop, one might need an irrigation system, but an individual cannot construct such a system without help. In addition, when there is enough density in the population to make it likely that deviations from norms will be observed and punished, there is more cultural tightness. The organizing theme is the importance of norms.

Loose cultures are usually exposed to several normative systems, so that people must develop tolerance for behavior that does not follow a single set of norms. Thailand is a culture that has been influenced by both the Indian and Chinese cultures, so it is understandable that it is loose. The organizing theme is the unimportance of norms.

1.3. Collectivism

This cultural pattern occurs in most of the traditional cultures of the world, including East Asia, Latin America, and Africa, as well as among Asian Americans. In collectivist cultures, the self is defined as an aspect of a group such as the family or tribe. In addition, religious, political, occupational, aesthetic, educational, and athletic groups can function as in-groups. In these cultures, the self has much social content (e.g., "I am a member of this family," "I am a member of this religious group") and is interdependent with one or more in-groups. When there is a difference between personal and in-group goals, in-group goals have priority. Behavior is a function of both attitudes (what one would like to do) and norms (what one should do). Group-based relationships endure even if the individual would like to get away from them. The organizing theme is the importance of the in-group.

1.4. Individualism

This cultural pattern occurs primarily in Northern and Western Europe as well as among European Americans, Australians, New Zealanders, and others influenced by Western culture. In this case, the self is

defined as autonomous, is independent of groups, and has little social context (e.g., "I am busy," "I am proud"). When there is a difference between personal and in-group goals, personal goals have priority. Behavior depends much more on attitudes than on norms. Group-based relationships are subjected to a cost–benefit analysis (e.g., "Do I get out of this relationship more than I have to put in?" [if so, stay in the relationship; if not, leave it]). The organizing theme is the importance of the individual.

Within any culture there is considerable variability concerning the presence of behaviors that correspond to collectivism or individualism. There are countercultural personalities in all cultures. People who tend toward individualism find collectivist cultures to be oppressive and so try to leave them. People who tend toward collectivism find individualist cultures to be cold and so join many groups (e.g., gangs, unions, communes).

1.4.1. Some Contrasts between Individualism and Collectivism

People in individualist cultures tend to see the self as stable and the environment as changeable (e.g., "If I do not like my job, I change jobs"). Conversely, people in collectivist cultures tend to see the environment as stable and themselves as changeable or ready to "fit in."

In individualist cultures, people are most likely to sample cues about events "inside" other people (e.g., beliefs, attitudes, values), whereas in collectivist cultures, they are most likely to sample cues about external events and social entities (e.g., norms, roles, situations, social structures, agreements, intergroup conflict).

In 1999, Norenzayan and colleagues reviewed evidence showing that when East Asians make dispositional attributions, they see traits as quite malleable, whereas when Western individualist samples make dispositional attributions, they see them as fixed. The authors reviewed a wide range of information, from laboratory studies to ethnographies, and concluded that probably all cultures make dispositional attributions but that cultural differences occur because people from East Asia make situational attributions much more frequently, and to a greater extent, than do people from the West.

When people communicate in individualist cultures, they sample the content most heavily, whereas when people communicate in collectivist cultures, they sample the context of the communication (e.g., level of voice, eye contact, gestures, emotional expression) most heavily.

Antecedents of individualism include affluence, migration (leaving the in-group), leadership roles, education, and living in heterogeneous environments (e.g., large cities). Antecedents of collectivism include poverty (survival depends on the help of the in-group), stable residence, low social class roles, and living in homogenous environments (e.g., rural or simpler societies).

1.4.2. Kinds of Individualism and Collectivism

These cultural patterns take different forms in different cultures. Specifically, there are horizontal and vertical individualist and collectivist cultures. Horizontal individualist cultures, such as Sweden, emphasize that the individual is independent from groups and is self-reliant but also that people do not wish to "stick out." Modesty is a virtue. People wanting to be "the best" characterize vertical individualist cultures such as the corporate and academic cultures in the United States. Competition is high and modesty is not a virtue is such cultures.

Horizontal collectivist cultures, such as the Israeli kibbutz, emphasize interdependence of the individual and the group, but there is little hierarchy. In theory, every member does all of the jobs, whether the jobs are prestigious or not. In vertical collectivist cultures, such as China and India, sacrifice for the group is virtuous. The individual is not important, and the group is all-important.

Another variety of collectivism is the one that stresses the relationship of children to all of the women in the extended family, as occurs in India, where there is no particular emphasis on the mother–child relationship. This pattern contrasts with the Japanese pattern, where the mother–child relationship is all-important and the relationship of the child to the women of the extended family is not important.

1.5. Hierarchy Versus Equality

This pattern is the same as the vertical–horizontal pattern just discussed, but it is useful to categorize it as a separate syndrome because it is very important and can be conceived independently from individualism–collectivism. In hierarchical cultures, people pay a lot of attention to the status of individuals when organizing their social behavior. In cultures that stress equality, people espouse many ideas, such as "one person/one vote," that stress the important of equality. The organizing themes are the importance or unimportance of status.

The syndromes that follow have not been studied well but are included here to suggest future research.

1.6. Masculinity Versus Femininity

In masculine cultures, men rarely do the jobs that women typically do. In feminine cultures, both men and women tend to do most of the tasks of the society (e.g., child rearing). The organizing theme is the importance or unimportance of being male.

1.7. Pragmatic Induction Versus Ideological Authoritarianism

This syndrome contrasts cultures in which people make most judgments based on information that comes from experience with cultures in which people make most judgments based on information obtained from ideological sources such as leaders, tribal chiefs, and gods. The organizing theme is empiricism versus ideology.

1.8. Short Versus Long Time Perspective

In cultures with a short time perspective, the emphasis is on the "here and now," whereas in cultures with a long time perspective, the emphasis can be on the past or future. The organizing theme is the present versus the past or future.

1.9. Valuing Planning Versus Spontaneity

In cultures that value planning, much effort is expended in planning, whereas in cultures that value spontaneity, planning is considered to be an imposition on one's freedom of action. The organizing themes are planning versus spontaneity.

1.10. Universalism Versus Particularism

In universalistic cultures, one uses general principles such as "to each according to his or her contribution," whereas in particularistic cultures, one takes into account who the other person is, such as in-group membership, previous accomplishments, sex, age, and social class, and one gives what is appropriate for a person who has those particular qualities. For example, in waiting for food at a cafeteria, universalistic cultures use the principle of "first come/first served," whereas particularistic cultures unquestionably allow high-status persons to go to the front of the line. The organizing principle is the use of universalistic criteria (one behaves in a certain way no matter who it is) versus particularistic criteria (behavior depends on who the other person is).

1.11. Specificity Versus Diffusion

Specific cultures tend to be analytical, paying attention to each element of information such as who said what, when, where, and why, whereas diffuse cultures tend to be holistic, assuming that if a positive (or negative) element is present in one element, it is also present in all of the other elements. The organizing principle is specificity of the stimuli versus the diffuse reaction to the stimuli.

1.12. Process Versus Outcome

In some cultures, people emphasize the process (e.g., "This is the way we do things"), whereas in other cultures, people pay little attention to the process but instead emphasize the outcome (e.g., "We must win"). Thus, in the former kind of culture, one pays attention to how something is done, whereas in the latter kind of culture, the emphasis is on what happens as a result of the action. The organizing principle is emphasis on process versus emphasis on outcomes.

1.13. Additional Syndromes

In addition to the syndromes just listed, there are many narrow syndromes. There is a literature, for instance, that describes syndromes whose organizing principles are honor, achievement, competition, and so on.

2. RELATIONSHIPS AMONG THE CULTURAL SYNDROMES

There is some evidence that syndromes are not unrelated to each other. For example, cultures that are tight tend to be collectivist. In general, collectivist cultures tend to be simpler and tighter, whereas individualist cultures tend to be more complex and looser. But there are many exceptions to these patterns. One usually finds more masculinity in collectivist cultures and more emphasis on pragmatic induction in individualist

cultures. Planning is more likely in individualist cultures. Particularism is more common in collectivist cultures. It is best to measure each cultural syndrome independently of the measurement of the other syndromes.

3. PRACTICAL SIGNIFICANCE

There is a correlation between some of the cultural syndromes and specific patterns of behavior. For example, industrial organizations in collectivist cultures are more likely to hire individuals who are in-group members regardless of their level of competence. In collectivist cultures, people tend to think of individuals as being easy to change and ready to adapt to different groups. Thus, it is understandable that they think that any in-group member can become a good employee. Loyalty to the organization, and especially to its authorities, is valued more in collectivist cultures than in individualist cultures. Thus, hiring loyal in-group members is more likely in collectivist cultures than in individualist cultures.

Collectivist managers are more likely to promote on the basis of tenure in the organization than on the basis of individual achievement. Also, collectivist employees prefer to be paid based on the achievement of the group rather than on the basis of individual achievement. But other things being equal, there is more organizational commitment and more training offered to employees in collectivist cultures that in individualist cultures.

Although on average members of a culture may have a particular attribute, there is very high variability within each culture. Thus, one can find employees who are behaving in highly individualistic ways in collectivist cultures and can find employees who are behaving in collectivist ways (e.g., joining unions) in individualist cultures. There are organizations that are individualist (e.g., academia) and organizations that are collectivist (e.g., the military) in all cultures.

When distributing resources such as bonuses, collectivists are likely to take into account the needs of employees (e.g., they have many children) or distribute equal amounts to each employee. In individualist cultures, bonuses go to those who have contributed more to their organizations.

The ideal leader is a benevolent father figure in collectivist cultures. The ideal leader is warm and supportive, but he or she also emphasizes production. Paternalism is a desirable leadership pattern in many collectivist cultures but not in individualist cultures. In fact, top management is much more aware of the private lives of employees in collectivist cultures than in individualist cultures. For example, top management in collectivist cultures likely will send condolences to an employee who had a death in the family (e.g., when an employee's parent dies), whereas top management in individualist cultures might not even be aware of the death of an employee's parent. Also, management in collectivist cultures likely will send congratulations when an employee's child graduates from high school or is accepted by a college.

Cultural differences in tightness also have implications. For example, people follow rules and specifications in manufacturing much more rigidly in tight cultures than in loose cultures. Improvisation and exploration of different ways in which to do the job are more common in loose cultures.

Differences in the way in which time is used provide a major difficulty in getting along with people from other cultures. Some cultures use monochronic time (e.g., holding one conversation at a time), whereas other cultures use polychronic time (e.g., holding several conversations with different people at the same time). When people from the former cultures are not aware of this cultural difference, they are easily offended by people from the latter cultures. Another contrast is between work time (i.e., time to work and not socialize) and social time. The definitions of what are work time and social time are culture bound. In some cultures, all of the farmers work together on the plot of one farmer and then go together to work on the farm of another farmer, and so on, and all the time they mix work time and social time. Still another contrast is between doing time (i.e., when one is supposed to do things) and not doing time. The desirability of "doing nothing" is very culture bound. In some cultures, it is the ultimate goal, whereas in other cultures, it is associated with utter boredom. Finally, the contrast between clock time (e.g., meetings end at a certain time) and event time (e.g., meetings end when the purpose of the meeting has been accomplished) can be the source of many misunderstandings.

See Also the Following Articles

Conformity across Cultures ■ Cultural Complexity ■ Cultural Psychology ■ Industrial/Organizational Psychology across Cultures ■ Intergroup Relations and Culture ■ Leadership and Culture ■ Motivation and Culture ■ Perception and Culture ■ Values and Culture

Further Reading

Carpenter, S. (2000). Effects of cultural tightness and collectivism on self-concept and causal attributions. *Cross-Cultural Research*, *34*, 38–56.

Chick, G. (1997). Cultural complexity: The concept and its measurement. *Cross-Cultural Research*, *31*, 275–307.

Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An "experimental ethnography." *Journal of Personality and Social Psychology*, *70*, 945–960.

Glenn, E. (1981). *Man and mankind: Conflicts and communication between cultures*. Norwood, NJ: Ablex.

Levine, R. (1997). *A geography of time*. New York: Basic Books.

Markus, H., & Kitayama, S. (1991). Culture and self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.

Norenzayan, A., Choi, I., & Nisbett, R. E. (1999). Eastern and Western perceptions of causality for social behavior: Lay theories about personalities and situations. In D. A. Prentice, & D. T. Miller (Eds.), *Cultural divides: Understanding and overcoming group conflict* (pp. 239–272). New York: Russell Sage.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.

Cyberpsychology

José M. Prieto

Complutense University, Madrid, Spain

GLOSSARY

cybernetics A discipline devoted to the study of information processing and communication systems as well as the functioning of control systems when applied to machines, animals, and organizations.

cyberpsychology A new realm of psychological research and action that uses cyberspace as an instrumental means toward an end, analyzes and identifies the underlying dimensions that structure and make sense in cyberspace, studies the cognitive and social processes that facilitate or impede the evolvement of cyberspace, discloses pragmatically designed principles leading to success in cyberspace, pays attention to personality disorders or social conflict arising in cyberspace, and introduces quality filters for searching and evaluating Internet resources.

cyberpunk A person who knows how to use computers as a means of self-expression and as a tool for producing virtual worlds in which reality and fantasy are intertwined and the individual person becomes a character in a scenario.

cyberspace An imaginary and multidimensional space in which people can absorb sensory experiences, enjoy intellectual creations, and interact cognitively and emotionally. It is also an economic survival kit for the information society.

cyborgs Initially, fusions of machine and organism, animate and inanimate at the same time. Later, new-age chimeras blurring mythic hybrids such as man/machine, male/female, white/black, and homosexual/heterosexual, created for pleasure by fine artists and graphic designers; fabulous creatures amalgamating over-elaborate features rooted in science fiction and social reality.

digital game-based learning A type of learning through use of a fusion of computer games and video games in which the game player is an actor exploiting the interactive possibilities, the climate of incertitude, and the curiosity appeal of a given game to attain learning goals.

e-learning The use of Internet and digital TV technologies to make available a broad array of information and databases that must be transformed in the learning process to produce sound and reliable knowledge and to facilitate human performance in real-life settings.

homepage Opening page of a Web presentation that provides links to other documents and that becomes the starting page to which secondary documents are linked. It is often the basic platform of a given Web-based or Web-linked organization.

hypertext Basic unit of associated sections within and outside a text connected to each other nonlinearly. Hypertext was invented by Douglas Engelbert in the mid-1960s, and it has become the underlying podium for the availability of documents in cyberspace.

knowledge management Active and systematic process of managing information to transform it into sound knowledge by capturing, creating, codifying, distributing, and filtering among networked users. Sometimes it is a substitute for the concept of learning.

telematics An English transposition of the French expression *telematique*, derived from the expression *teleinformatique*, that may be translated as "networked computers" or "wired computers."

virtual reality The computerized simulation of an entire three-dimensional surrounding in which a person can perform actions or manipulate objects by clicking on the mouse. Visually the degree of realism is very high but kinetically leaves a lot to be desired.

World Wide Web (Web) A hypertext-based information network rooted in internationally accepted standards. It was developed in 1990 at the European Centre for Particle Physics by Tim Berners Lee.

This article has been divided into three sections. The first analyzes the emergence of cyberspace as a counter culture, emphasizing distinctive backgrounds. The second analyzes how cyberpsychology is a brand new name in the phase of increasing its visibility within the domain of applied psychology. The third pays attention to six different strategies used by applied psychologists when they deal with Internet and multimedia databases and hyperdocuments.

1. CYBERSPACE: A TECHNOCULTURAL REALM

There is a hidden nexus between the basic unit of electric current, known as the ampere, and "cyber," a multifaceted root present in a large variety of domains within the umbrella of new information and communication technologies (NICT). The term ampere expresses homage to André Marie Ampère (1775–1836), a French physicist, mathematician, and philosopher. He coined the term cybernetics in 1834 for a prospective new science devoted to the overseeing and control of governments. In classic Greek, $\kappa\upsilon\beta\epsilon\rho\nu\acute{\alpha}\omega$ means to lead or to pilot, and so the derived noun, $\kappa\upsilon\beta\epsilon\rho\nu\eta\tau\iota\kappa\acute{o}\zeta$ means the pilot. Nobert Wiener (1894–1964) recaptured the word cybernetics in 1948 as a title for his book on "the science of control and communication in the animal and the machine." Later, in 1950, he emphasized that cybernetics should ultimately be regarded as a matter of "the human use of human beings," highlighting the nexus between learning and feedback as both technological and cultural frameworks in the functioning of control and governing systems.

During the 1950s and 1960s, cybernetics became a widely accepted label for studying (1) the processing of information and communication systems, (2) the flow of information within a system, and (3) the use of feedback to get intended impacts on goal-directed activities cropping up in technological artifacts as well as in living organisms and organizational settings. Soon cybernetics lent its name and the new root "cyber" started to circulate, becoming used by computer-literate university graduates in a large variety of disciplines.

In 1970, Alvin Toffler coined the term "cyborgs" in his book *Future Shock*, which devoted several pages to analyzing the possibilities of human–machine integration and the interaction of human brains and databases through networked communication. Initially, a cyborg was described as a fusion of machine and organism, animate and inanimate at the same time. Mechanical or electronic devices built into the body allowed its physiological functioning to extend some abilities beyond normal or to compensate for some disabilities. Psychological expertise was demanded during the design process to optimize performance levels or to smooth training processes and after the surgical implant of devices to facilitate the psychological adjustment of ongoing interactions.

During the 1980s, the conception of cyborgs transformed classical dichotomies into mythic hybrids by blurring binary pairs such as man/machine, male/female, white/black, and homosexual/heterosexual. At first glance, it was a change of focus: cyborgs started to be created for pleasure. Fine artists and graphic designers crossed the line of gender representations and pushed their art into the awareness of the sexual impact of cyberbodies. So, cyborgs became sexy, and a diversified fauna of new-age chimeras emerged. "A cyborg is a cybernetic organism, a creature of science fiction and a creature of social reality," highlighted Haraway (1991) in her seminal essay "Cyborg Manifesto" published in 1985. From a psychological perspective, these utopian cyberbodies induced morbid fascination and were ready for scrutiny as the latest taste in projective techniques. A cyborg is a cybernetic organism, a creature of science fiction and a creature of social reality.

The term "cyberpunk" started to be used as an argot among writers in counter-cultural circles in the mid-1970s, for example, by John Brunner in his novel *Shockwave Rider* and by William Gibson in several science fiction stories collected in *Neuromancer*. Soon, some creative writers began to call themselves "cyberpunks," stressing that they used computers and video games as a means of self-expression, personal pleasure, net profit, and out of a sense of duty. In fact, they ascribed the roles of main or secondary characters to "first-generation cyborgs," that is, microsurgically reworked characters, showing a disembodied

intelligence, bringing to the fore an anarchist cultural background melting away into a high-tech mood and open-minded attitude. In cyberpunk stories, the narrative usually took place in postindustrial and information-governed settings populated by urban misfits, dead reconstructions of people previously alive or artificially smart beings ready to think for themselves and question authority. Cyberpunks were often depicted as skillful people with expertise in knowing how to take advantage of NICT to attain goals that were often malevolent.

During the 1980s, fiction writers realized that they had developed a set of beliefs about the existence of some kind of actual space behind the screen, "a place that you cannot see but you know is there." This world of lucid dreaming beyond a television or a computer screen came to be known as cyberspace, an imaginary and fictional universe in which sensory experiences take place, the mind is absorbed, and the person feels as one with the set of stimuli, challenges, and performances elicited. It generates a trancelike experience, actual and genuine, known as "consensual hallucination," the assertion coined by cyberpunk writers. Cyberspace was a place where individuals were welcome if they knew how to stroll along boundless passageways, entering and exiting large intellectual creations made available online. It was a "World III" structure in the terminology launched by Sir Karl Popper (1902–1994).

Cybersex was a term introduced during the 1990s to identify the existence of sensual and hedonistic waterways irrigating the digital underground of cyberspace. It is the domain of a large variety of incorporeal intercourses, sometimes known as "robocopulation" or "onanism for two," that make the most of computer-generated hyperrealism as well as mind-to-mind networked but yet dreamed scenarios. As Dery put it, "sex with machines, together with dalliances conducted in virtual worlds, seems a seductive alternative in an age of AIDS, unwanted pregnancies and sexually transmitted diseases" (1996, p. 199). It is a kind of online sex play, X-rated and interactive computer programs adapted for the stage by similarly equipped participants recurring to an array of sensor-effectors ready to generate a realistic sense of tactile presence.

Bell and Kennedy and Spiller have compiled large sets of difficult-to-find articles on cyber-cultural challenges and dilemmas, published over several decades during the second half of the 20th century, that otherwise would be rather inaccessible.

2. CYBERPSYCHOLOGY: AN INCREASINGLY GROUNDED BRAND NAME

The term cyberpsychology started to circulate in cyberspace in 1994 when Prof. Leon James at the University of Hawaii wrote a pioneering paper, "Cyberpsychology: Principles of Creating Virtual Presence," which was expanded over a decade and made available online. As a result, some post-modern views on psychology and Jungian connotations were launched: an ineffable continuity of an online communal mind exists and grows in cyberspace as a self-ruling collective unconscious. The existence of a critical mass in interactions and transactions "in any topical zone and the type of ongoing activity" suggest the emergence of a communal mind in the Internet. "Topics and activities in cyberspace create their own virtual zones that become accessible to others across time and space" and give the impression of omnipresence in such a communal mind. The overlap between cyberspace and the mind was asserted under the umbrella of cognitive psychology. Computers were appraised as "convenient and powerful extensions of the human mind" and the consequence was that "every characteristic of the mind can be expected to show up as a property of the cyberspace." The basic argument justified a new denomination such as cyberpsychology based on the certitude that "cyberspace, like mind, is not in physical space, but in virtual space, without extension, distance, or mass."

In 1994, José M. Prieto at the University of Madrid set up the homepage of the International Association of Applied Psychology (http://www.iaapsy.org), which included the online proceedings of the 23rd International Congress of Applied Psychology. Nielsen, the leading expert in usability challenges, advances, and dilemmas in the production of multimedia and hypertext documents, mentioned this initiative as an illustration of navigating large information spaces. Soon afterward, Prieto began to focus on cyberspace as a setting where psychological teaching, learning, and research programs may take place and where spatial orientation and time management skills play a very significant role in online performance. The initial focus was outlined in the paper "Psychology and Telematics" (a new term derived from the combination of "tele" and "informatics"), and the next step was the interface between literacy in NICT and the involvement of work and organizational psychologists in knowledge management programs. In 1997, it became an e-learning space

nicknamed *prietolandia* by psychology students: an online learning experience for university graduates in psychology from 1995 to 2003, when the program came to an end.

In 1996, Prof. John Suler at Rider University started to write an electronic book, *The Basic Psychological Qualities of Cyberspace*, made available online, another project in a course of continuous expansion. The point of departure was an intensive case study carried out on the psychological and social interactions held by members of a very vivid and evolving community who used to meet in a quiet notorious virtual space known as "The Palace" (http://www.thepalace.com) that closed by the end of the 1990s but was again active by 2004. In 1995–1996 it was a visual, auditory, and chat environment where visitors adopted and role-played a large variety of characters by using graphical representations in the realm of pure fantasy and fiction. Suler used the participant observation technique to study cyberspace as "a psychological space" and as "a dream world." The psychoanalytic as well as field theory perspectives shaped the theoretical and operational background, making sense of what emerged in this research project. Internet users felt comfortable because somehow cyberspace was "an extension of their mind and personality—a 'space' that reflects their tastes, attitudes, and interests." Following a psychoanalytical vein, cyberspace was interpreted as a "transitional space" between the self and the other.

A peer-reviewed journal entitled *Cyberpsychology and Behavior* was launched in 1998 (http://www.liebertpub.com) and made public during the 24th International Congress of Applied Psychology and the 106th Annual APA Convention, both held in San Francisco in August 1998. Since the year 2000, the electronic and the printed version have been accessible at the same time; subscribers have access to each online issue via a password.

Gackenbach compiled 13 contributions written by 20 authors from Australia, Canada, Israel, the United Kingdom, and the United States focusing on the psychology of Internet use. Special attention was paid to the normal and deviant aspects of the self when staying online, lively relationships maintained during close and distant interactions, and transpersonal issues such as the development of consciousness and collective unconscious. Through surveys and case studies, the incidence of Internet addiction was examined and its presence was restricted to "an exceedingly tiny minority" (p. 71). This is a typical example of research findings that refute comments in the mass

media occasionally introduced by practitioners expressing opinions without reliable data on the subject.

Wallace approached the Internet within a psychological perspective; the outcome was a book titled *The Psychology of the Internet*. It focused on the "online persona" and the way Internet users generate impressions, use masks as ID during interactions, and become involved in group dynamics through conformity and cooperation as well as conflicts and aggressions. She studied how liking and loving take place on the net, calling attention to how interpersonal attractions arise when people try to find out who is next door on the net.

Fink is the lead author of a book devoted to analyzing how to use computers and cyberspace in the clinical practice and psychotherapy. There are 10 separate contributions that each highlight specific topics. Together, they scrutinize the challenges faced by professional psychologists determined to transmute the conventional couch into an online couch, to explore the self by inviting the patient to produce a personal Web page, to go after the consequences of a disembodied gender in cyberspace, and to go along with the consequences of virtual healthcare programs. A catalogue of Internet resources for clinical psychologists was added as a bonus.

Gordo-López and Parker, in a book entitled *Cyberpsychology,* brought together a set of 14 articles written by 16 psychologists and social science scholars from Europe, the United States, and Latin America, bringing into focus the relationships between psychology and cybernetics. The prevailing approach is rather theoretical, rooting psychology into the technocultural background of cyberspace without endorsing fantasies of liberation that surround NICT in the mass media and the popular culture regarding the Internet. The approach is rather interpretative: "Cyberpsychology needs to embed within itself a self-annihilating device, a certain kind of critical and self-critical narrative"; "We want to insist that it is only in its ephemeral use that the potential of cyberpsychological critique can function" (p. 16).

Birnbaum compiled 12 contributions produced by 23 authors from Canada, Germany, Italy, Switzerland, the United Kingdom, and the United States. Studies carried out in the Internet were compared to similar studies carried out in laboratories or in field studies, stressing methodological aspects to be taken into consideration when deciding which is the appropriate medium to use. The role of individual differences and cross-cultural differences as identified through

online research was also examined. Potential uses of advanced computer techniques or pondering the appropriateness of server-side solutions reveal weak points in design as well as the adequacy of server–client relationships. This is an Achilles heel rarely assessed in papers.

Wolfe compiled 12 contributions written by 19 authors from Europe and the United States that examine cost-effective advances and achievements attained by the use of learning technologies placing the development of Web resources on a consistent psychological foundation. It combines theoretical advances and empirical findings and moves from the frame of reference of the individual learner to that of learning communities in university campuses. One out of two college courses list Web resources in the syllabus, and one out of four college courses have their own Web pages, thus producing a gap between those making an offer of online and offline handouts or practical exercises, for instance.

Prensky has emphasized the importance of computerized games as educational and training tools that meet the needs and learning styles of people wishing to improve their expertise by using simulations and theatrical scenarios. There is in fact a generational gap between educators belonging to a pre-digital cohort and students raised in an audiovisual and digitalized cultural background. Digital game-based strategies allow the acquisition of very specific skills on demand for specialized jobs and risky situations. Findings and developments founded on studies derived from the psychology of play provide the background for the cognitive and emotional involvement of students galvanized as players.

Reips and Bosnjak compiled 20 contributions written by 40 authors from Europe and the United States. These papers obtained expert advice and commentaries from two rounds of double peer review from 32 independent reviewers, evidencing high standards in standard publication procedures. They examine the Internet both as an instrument for, and an object of, scientific research, and approach issues such as psychological Web experiments, Web questionnaire studies, studying perception on the net, net-based surveys, and communication research as well as knowledge acquisition and learning throughout cyberspace.

Stein has written an essential Baedeker for students and scholars, a traveler's guide to the topography of cyberpsychology. His book is a comprehensive and accessible overview of what resources are available, how to track down subject-specific materials and use them efficiently, and how to compile Internet references for term papers, reports, and dissertations. The author has also provided an online catalogue of appropriate hyperlinks mentioned in the book.

3. CYBERPSYCHOLOGY: SIX PERSPECTIVES AS OPERATIONAL TOOLS

It is possible to distinguish six quite different strategic approaches followed by applied psychologists when they use the Internet and multimedia hyper-documents in daily professional activities. Private life approaches are not considered here, but they are always embedded as background. The six professional strategies advanced are (1) instrumental, (2) dimensional, (3) process-based, (4) pragmatic, (5) psychopathologically biased, and (6) knowledge management-based.

3.1. A Tool for Psychologists

The first strategy approaches online computer-based resources and support facilities as a tool for conducting a large number of activities that psychologists carry out regularly in university campuses as well as in professional practice. It is an instrumental way of understanding psychology within the cyberspace, the nexus "means to an end." Cyberspace is viewed as a means to (1) communicate via email or transmit documents, (2) participate in discussions via interactive groups and Web-based conferences, (3) exchange or supply files and programs, (4) carry out transactions with customers, and (5) increase visibility across borders. Thus, cyberspace is valued as a useful territory that facilitates the achievement of certain ad hoc aims, rooted in psychological action and effective performance. Psychology students and graduates typically learn to handle only a few of the tools available, those they become comfortable with and use for certain purposes. Newer tools are often avoided because they demand too much time and imply too much incertitude: only members of younger cohorts are open minded enough to try them. This attitude contrasts with the hedonistic ways of understanding cyberspace: a stage for playfulness and inventiveness, a joyful scenario for divergent thinking among advance expert

knowledge-based groups ready to fix acceptable standards in 5 or 10 years.

also the basis for the study of human diversity in cyberspace.

3.2. Alternate Dimensions of Experience

The second strategy deals with psychological dimensions of computer-based settings, online environments, and personal and professional grounds of performance instigated by the involvement in computer networks. It is the realm of magnitudes and measurements present in online phenomena and online users. It brings to the forefront the psychological analysis of personality questionnaires, for instance, producing comparisons between factor structures obtained through offline and online versions of the same instrument. Present evidence suggests equivalence regarding what they measure and how well they serve the intended purpose. Similar approaches have been tried in the study of abilities, skills, interests, and values underlying online tasks and activities examined by researchers to obtain the factors or clusters underlying effective or ineffective performance. It allows the identification of multifaceted features such as experience requirements, cyber-users' characteristics and requirements, and communalities and specificities among members of relevant virtual communities. Follow-up studies of observable response patterns obtained through Web-based surveys favor the reconstruction of response undertakings and the identification of specific typological profiles different from those obtained via offline surveys. For example, controlling the display of questions and answering actions allows the number of questions displayed and the number of answers marked to be taken into consideration. Findings support the distinction between "performers" and "explorers" as basic typology. Performers answer all questions displayed, whereas explorers read initial set of instructions and preliminary items and quit suddenly, read the full set of questions screen by screen but never answer, or answer a few or even many questions but do not conclude the protocol. The psychological study of group structures in cyberspace also follows a dimensional approach, since it allows comparisons within and between communities to identify overt or hidden fraternities in communication networks, leadership styles in online chats, and prevailing norms and roles among users when they are on stage in a Web conference room, for example. Psychological dimensions are

3.3. Psychological Processes and Human Performance

The third strategy addresses the issue of the psychological processes associated with human performance in networked computers and hypermedia. This process-based notion of cyberpsychology stresses that analysis may be carried out at the individual or at the group level and that it requires follow-up controls and strategies. The distinction between text-based and hypertext-based handouts used in online seminars has served to analyze operational and cognitive interactions concerning navigation, information scrutiny, and knowledge acquisition. Feelings of getting lost while navigating hypertext documents have been detected, and organization and presentation schemes have been suggested and validated to provide reliable frameworks for interested readers. For instance, the use of frames that include a contents index or a graphic map synchronized to sequential actions seems to facilitate the orientation of individual readers. Generational gaps have been remarked upon, and the presence or absence of dynamic icons accentuates the impression of "cool" versus "puerile" settings and thus acceptance or rebuff, by cohorts, of information presented. Eye movement screenings allow the analysis of how much time readers stay absorbed in a given text paying or not paying attention to navigation bar as compared to readers looking through a hypertext document. The navigation process seems to distract persons who are not experts in the subject from the actual text; they retrieve more information compared to those who are expert in the subject who can skim the text with hypertext advantages. Sensation, perception, attention, memory, knowledge acquisition, conscious experience, tacit learning and so on play an active role in the way people behave in cyberspace and have been examined, capitalizing on models and procedures developed in cognitive psychology. The study of group processes in cyberspace enables the discernment of formation, commitment, and identity within a given virtual community, as well as the understanding of influences pervading each community, of decision-making patterns and the value of minority dissent, of conscious and unconscious phenomena, and of autonomy and effectiveness.

3.4. Optimizing the Cyberspace Experience

The fourth strategy tackles the understanding of useful principles and guidelines based on psychological grounds that may improve the success of a home page, a portal, or an online e-business or e-commerce action plan. It is a pragmatic notion of cyberpsychology that looks for psychological theories, models, or findings that may refine communication patterns, learning strategies, and the intended presence and visibility of individuals and groups in cyberspace. The focus of attention is the identification of how cyberspace can be used to isolate and advance a psychological elucidation of behaviors and actions and the way of influencing its appearance or withdrawal. For instance, Gestalt psychology principles, formulated at the turn of the 20th century, are still an adequate framework in which to ameliorate the design of multimedia pages and navigation tools. Tacit learning theories have been implemented to reduce the number of procedural steps in online commercial or stock market transactions. Heuristic techniques have been used to optimize the testing of complex hypermedia documents to identify missing links, errors, intended loops, or unfortunate paths, before producing the final version. The pragmatic approach favors a focus on online behavior as natural, on online inventiveness in problem-solving strategies as natural, and on accepting a continuous flow of new online standards and norms as natural.

3.5. Cyber Pathology

The fifth strategy focuses on online psychological disorders; online overuse is the recurring example. In fact, it is the consequence of a parody written in 1995 by Ivan Goldberg, a psychiatrist who is quite active online, that made fun of the pathological gambling criteria described in DSM-IV. This resulted in a psychopathologically biased notion of cyberpsychology pivoting on reported cases of pornography, interpersonal or sexual attraction, disembodied gender, marital problems, neglect of children, and so on, as broadcast by the mass media. The high visibility of this perspective resulted in focus on antisocial and manipulative behavior online. However, it must be stated that cyberspace just mirrors a person's regular social behavior and is simply a new scenario for interrelating with others. There is no evidence suggesting that cyberspace may be the direct cause of new mental or emotional problems. It is more a matter of inappropriate virtual behaviors such as (1) spam, invasive, unsolicited, and superabundant email, (2) "spoof," a message sent with the sender's identity concealed, and (3) "flaming," communication that is perceived as an insult, a direct attack, abuse, insolence, or rudeness. The large majority of cyberspace dwellers understand the consequence of spamming or spoofing because their email inboxes regularly receive a large number of unwanted or forged email messages. These behaviors have highly detrimental consequences, since many derived costs are borne by the person receiving the messages instead of the sender. Paradoxically, these inappropriate behaviors can shock online users, but usually are not considered scandalous or outrageous by offline users such as investigative journalists or psychologists in the field. This is just one example of double standards regarding criminal behavior in cyberspace versus in "real life."

3.6. Information and Knowledge Management

The sixth strategy views cyberspace as a huge virtual library that facilitates access to information in accounts, advertisements, books, brochures, databases, discussion groups, documents, journals, manuals, newsgroups, papers, journals, newsgroups, proceedings, software, surveys, syllabi, and so on. At first glance, cyberspace can be viewed as a storehouse ready to grasp the interest of psychologists interested in keeping an open mind to a large variety of old and new subjects, theoretical, technical, or practical. However, the use of filters is required to separate the wheat from the chaff. The same criteria that are useful in the analysis of scientific papers can be used to identify helpful online resources: (1) intended audience, (2) authority on the subject, (3) amount and type of coverage (e.g., in-depth versus superficial), (4) milieu in which the information is accessible, (5) objectivity versus subjectivity in distinguishing facts from values or opinions, (6) precision according to reliable sources, (7) topicality (i.e., high or low profile subjects) and, (8) the ease of use of the interface.

See Also the Following Articles

Internet Counseling

Further Reading

Bell, D., & Kennedy, B. M. (2000). *The cyberculture reader*. London: Routledge.

Birhnhaum, M. H. (2000). *Psychological experiments on the Internet*. San Diego, CA: Academic Press.

Dery, M. (1996). *Escape velocity: Cyberculture at the end of the century.* London: Hodder & Stoughton.

Fink, J. (1999). *How to use computers and cyberspace in the clinical practice of psychotherapy.* Northvale, NJ: Jason Aronson.

Gackenbach, J. (1998). *Psychology and the Internet: Intrapersonal, interpersonal and transpersonal implications.* San Diego, CA: Academic Press.

Gordo-López, A. J., & Parker, J. (1999). *Cyberpsychology.* London: Macmillan Press.

Haraway, D. (1991). *Simians, cyborgs and women: The reinvention of nature.* New York: Routledge.

Nilsen, J. (1994). *Multimedia and hypertext: The Internet and beyond.* Cambridge, MA: Academic Press Professional.

Prensky, M. (2001). *Digital game-based learning.* New York: McGraw Hill.

Prieto, J. M., & Sanz, P. (2002). Cyberpsychology. *In Encyclopedia of life support systems.* Oxford, UK: EOLSS. http://www.eolss.net.

Reips, U. D., & Bosnjak, M. (2001). *Dimensions of Internet science.* Lengerich, Germany: Pabst Science Publishers.

Spiller, N. (2002). *Cyber-reader: Critical writings for the digital era.* London: Phaidon.

Stein, S. (2003). *Psychology on the Web: A student guide.* Essex, UK: Pearson Education.

Wallace, P. (1999). *The psychology of the Internet.* Cambridge, UK: Cambridge University Press.

Whittle, D. B. (1997). *Cyberspace: The human dimension.* New York: W. H. Freeman.

Wiener, N. (1948). *Cybernetics: Control and communication in animal and machine.* Cambridge, MA: MIT Press.

Wolfe, C. R. (2000). *Learning and teaching on the World Wide Web.* San Diego, CA: Academic Press.

Decision Making

Helmut Jungermann

Technical University of Berlin, Berlin, Germany

GLOSSARY

biases Systematic judgmental tendencies that violate the rules of probability.

bounded rationality Decision making within the limits of human cognitive capacities.

choice The selection of an option from a set of options.

decision The commitment to a course of action.

heuristics Mental rules-of-thumb used for the judgment of probabilities.

image theory A theory that includes the maximizing strategy as a special case but focuses on a strategy that works on the compatibility of options with the decision maker's goals and plans.

prospect theory The currently dominant theory of individual decision making, developed by Kahneman and Tversky, that keeps the core components of the classical rational model but takes the findings of empirical research into account.

rationality Decision making consistent with the axioms of the normative model that assumes the use of all available information, stable preferences, and a maximizing strategy.

uncertainty/probability The uncertainty that people associate with the occurrence of the potential outcomes of their decisions.

value/utility The value that people attach to outcomes resulting from their decisions.

Decision making is a mental or behavioral commitment to a course of action. In a broader sense, the term "decision making" denotes an information-processing activity of a single decision maker, or of multiple decision makers, that begins with the recognition of a choice situation and ends with the implementation of the choice and the monitoring of its effects.

1. INTRODUCTION

More than many other behaviors, individual decision making is a subject not only of psychological research in the laboratory but also of interest in a number of "real-world" domains where significant decisions are actually made and where psychological knowledge has been applied successfully. The study of decision making has its roots in economics and statistics, and this is still the most active field for decision researchers. Topics include how people choose among consumption goods, investment options, and/or retirement plans as well as how managers make production and

marketing decisions and how they make organizational and employment decisions. Another important field involves health-related and medical decisions of both patients and physicians. For patients, topics include how people make their personal decisions about smoking, drinking, taking drugs, and/or sexual behaviors as well as how patients choose among treatment options (e.g., surgery or radiation) and how they can be helped in making such decisions. For physicians, topics include how they choose among diagnostic and treatment alternatives and how their individual or collective decision processes can be supported. A very different field is aviation, where the interest focuses on the decisions that pilots must make—often under extreme time pressure. Closely related are decision situations in high-risk facilities such as nuclear power and chemical plants. Knowledge about people's decision-making behavior has also been used for designing counseling and advice-giving procedures. Examples include when patients are asked for their informed consent, when parents come to a genetic counselor, when people seek investment advice, and when adolescents need behavioral strategies against sexual assault.

2. KEY CONCEPTS

A decision is a mental or behavioral commitment to a course of action. An option to change the status quo may be brought to the attention of a person and is accepted or rejected, or one of several given options is preferred over the others. A decision can be expressed by verbal judgment (e.g., "Option X is the best one") or by behavioral choice (e.g., Option X is chosen). In a narrow sense, the term "decision" denotes only the moment in which the commitment is made. In a broader sense, the term denotes an information-processing activity, usually beginning with the recognition of a choice situation and ending with the monitoring of the outcomes of the chosen option. Research has looked at all phases of a decision-making process—how people identify, generate, screen, and modify options; how they mentally simulate courses of action; how they match actions to situations; how they frame, compare, and evaluate all or a subset of options; and how they process information after a choice has been made. Decisions can require different amounts of cognitive effort and be made on different levels of awareness, depending largely on the significance of the problem, the familiarity of the situation,

and the experience of the decision maker. When people drive from their home to their office, or when surgeons perform appendectomies, their decision making is often nearly automatic and is experienced as performed routinely or intuitively. On the other hand, when people in a restaurant choose from a menu, they realize that they must make decisions and often ponder about the options and consider the various pros and cons. Typically, substantial cognitive effort is invested only if the problem is important, complex, or unfamiliar such as in the purchase of a house by a family, the choice of a cancer treatment by a patient, or the development of a new product by a company. Economists tend to consider a decision maker—the homo economicus—as a rational decision maker if he or she uses all available information, if preferences and beliefs are stable and consistent, and if a maximizing strategy is applied. Psychologists (and behavioral economists) favor the assumption that decision makers usually operate with bounded rationality (i.e., within the limits of their cognitive capacity) and that they apply the strategy that best matches their situational assessment and motivation.

3. FEATURES AND DISTINCTIONS

The dominant view is that people make their decisions as a result of anticipating and evaluating the potential outcomes and the uncertainties associated with them. This is called a consequentialist perspective. The major part of this article focuses on research that takes this perspective, but nonconsequentialist approaches are also presented later. From the consequentialist perspective, decision problems differ formally in the following ways. First, the consequences of the decision are certain (e.g., the consequences of choosing Apartment A or Apartment B) or uncertain (e.g., the consequences of choosing chemical therapy or surgery). Second, the consequences are unidimensional (e.g., the consequences of gambling in a casino as either a gain or a loss of money) or multidimensional (e.g., the consequences of choosing a job in a company or at a university). Third, the decision process may have one stage (e.g., the decision to hire a person) or several stages (e.g., the decision to develop a new drug). Fourth, the decision may be unique (e.g., decisions among cancer treatments) or may be made repeatedly (e.g., decisions between job applicants). Fifth, another important distinction must be made between individual decision situations and social

ones. In the former case, decisions are made by individuals, and individuals primarily must bear the consequences. This situation is the topic of cognitive psychology. In the latter case, decisions are also made by individuals, but the decisions are made in direct or indirect interaction among these individuals. This situation is the topic of social and organizational psychology.

4. VALUE/UTILITY

People are assumed to make their decisions depending on their preferences for the potential consequences (i.e., states or events) of the given options. Implicitly or explicitly, people evaluate these consequences, and the subjective values, called utilities, determine their decisions. However, the utility of outcomes is not simply a linear function of their quantity. In general, one finds a logarithmic function well known from psychophysics. In economics, this is called the decreasing marginal utility function: The second million causes less pleasure than the first million. The same holds for the negative part of the function: A dessert for $5 appears to be less expensive after a $70 dinner than after a $7 pizza. Advertising and marketing use the knowledge about the psychophysics of spending in many ways.

Evaluations are influenced by a number of factors, including the following prominent ones. First, people value an object more highly when they believe to have achieved it by their performance rather than by luck (i.e., source dependency). Second, people value an object more highly when they consider selling it than when they consider buying the same object (i.e., endowment effect); that is, a person would want a higher price for a painting he or she owns than the person would pay for the same painting. Third, the evaluation of an option is sometimes influenced by investment of money, time, or effort in the past (i.e., sunk cost effect); that is, a person does not leave a boring movie because he or she has paid for the admission ticket. Fourth, the evaluation also depends on the moment in time when the object will be received or when the state will occur (i.e., time preferences); that is, a person prefers to receive $100 sooner rather than later, and the person probably prefers to see the dentist later rather than sooner (unless he or she has pain).

Multiattribute options (e.g., consumption goods, job applicants) require trade-offs between attributes. Consumer reports provide the relevant attributes of the goods and the values of the goods on these attributes. But consumers must make the trade-offs themselves.

Numerous studies have examined which attributes people use in such situations, how they weigh them, and how they integrate the information about the goods on all attributes. Sometimes, people use noncompensatory rules; that is, they eliminate all options that do not exceed certain cutoff points on one or several attributes or that do not share certain features, independent on their values on other attributes. For instance, a person looking for an apartment may first screen out those apartments that do not have a minimum size, then those whose rents exceed a certain limit, then those whose distance to the beach exceeds a certain distance, and so on until finally one acceptable apartment is left. Alternatively, people may use compensatory rules; that is, they accept that each value on an attribute, however low, can be compensated by high values on other attributes. For the person looking for an apartment, the small size of an apartment might not lead to its elimination because it can be compensated by a low rent. Consumer research has used the theoretical concepts, the research methods, and the empirical findings extensively.

5. UNCERTAINTY/PROBABILITY

Decisions are always made under uncertainty because their consequences occur in the future and, thus, depend not only on the decision makers' choice but also on events and states not under the decision makers' control. In some situations, the consequences can be considered as nearly certain or the fundamental uncertainty is irrelevant for the choice. However, in most decision situations, the uncertainty cannot be ignored; therefore, research on how people deal with uncertainty is an important field of decision research.

For mathematicians, probability is the language of uncertainty; for decision researchers, subjective probability is the language of uncertainty. The subjective probability for an event may be based on the known or assumed relative frequency of that event (e.g., the subjective probability of rain tomorrow, the subjective probability of having a car accident), but in many situations the event is unique and there are no relative frequencies (e.g., the subjective probability of Brazil winning the soccer world championship, the subjective probability of an increase of the gold price by the end of the year). Research has studied the psychological variants of uncertainty, their cognitive characteristics and determinants, the ways in which people express uncertainty (e.g., numerically, verbally), and the ways

in which people understand expressions of uncertainty (e.g., probabilities, percentages, frequencies). Intuitive judgments of probabilities have been shown, under certain conditions, to deviate considerably from our statistical knowledge and to violate the rules of probability theory systematically. These observations led Tversky and Kahneman to develop their heuristics and biases research program. The key assumption is that people intuitively apply mental heuristics rather than statistical algorithms, and although these heuristics are generally useful and efficient, they can produce judgmental errors, fallacies, or biases. Two heuristics are particularly important: representativeness and availability. Representativeness is when events are unique (e.g., the guilt of a defendant) or set in the future (e.g., the outcome of the next election) and people draw on their knowledge about the case and judge the probability by the closeness of the match between the event and the prediction derived from their knowledge. Availability is when events are grouped in categories or by features (e.g., percentages of road accidents in a year) and people judge the probability by the number of examples that come to mind or by the ease with which they come to mind. Many judgmental biases can be explained by the assumption that people use these heuristics, for instance, the base rate fallacy, sample size neglect, regression neglect, the unpacking effect, and the conjunction fallacy. Other prominent biases in subjective probability judgments, not directly related to the use of heuristics, are the hindsight effect, the overconfidence effect, conversion errors, the illusion of control, and unrealistic optimism.

6. MODELS AND MODES OF DECISION MAKING

Prescriptive theories of decision making say how decisions should be made if the decision maker wants to follow certain principles of rationality (e.g., the principle of transitivity). These theories are part of management science. They provide formal rules and procedures for structuring and processing the relevant information and, thus, provide support in complex decision situations. Descriptive theories aim to show how people actually make decisions. Because the cognitive capacity is limited, or at least is not always used efficiently, actual decisions are often suboptimal compared with the decisions prescribed by formal theories.

The first descriptive model of decision making under uncertainty was proposed by Edwards in 1954. The subjective equivalent utility (SEU) model assumes that people try to maximize their subjectively expected utility. The SEU of an option is the sum of the utilities of its consequences, weighted by the subjective probabilities of their occurrence. The decision maker is assumed to choose the option with the highest SEU value. In 1979, Kahneman and Tversky proposed prospect theory (PT) as an updated version of the SEU model. The theory sticks to the assumption that decisions are determined by the values and probabilities of their consequences but takes into account the many observations of decision-making behaviors that do not concur with the SEU model. PT has become very influential, particularly in the economic and medical domain. In 2002, Kahneman was awarded the Nobel Prize in economics, primarily for this theory (Tversky had died in 1996). The essential elements of PT are the value function and the decision weight function. With respect to the value factor, PT assumes that people mentally code the potential outcomes of options in relation to a reference point, that is, the status quo or the aspiration level. Outcomes above the reference point are coded as gains and outcomes below the reference point are coded as losses. Gains and losses are evaluated according to a value function that has two properties. First, the function is concave over gains and convex over losses; that is, additional gains please less, and additional losses hurt less. Second, the function is steeper over losses than over gains; that is, a loss of $100 hurts more than a gain of $100 pleases. With respect to the probability factor, PT assumes that subjective probabilities of outcomes are transformed into weights that represent the significance of the occurrence of the outcome. For instance, many people weigh the transition from a 0.99 probability to certainty (1.00) as heavier than a transition from a 0.41 probability to a 0.42 probability. PT provides explanations for a number of empirical phenomena that often are considered as irrationalities or anomalies, particularly in economics. Examples include the finding that preferences are not invariant with respect to their verbal description (i.e., framing effect) and the finding that investors sell winner stocks too fast and hold loser stocks too long (i.e., disposition effect).

An alternative theory, image theory, was proposed in 1990 by Beach. Image theory has become very popular in management and business schools. Beach assumed that people apply an optimizing strategy like the one proposed in PT, only under very special conditions. More often, people examine whether a new option is compatible with their goals and plans (called images)

and accept and implement the option if that is the case. If the option is not compatible and violates important features, other options are searched and explored. When people are confronted with a number of options, they first screen out the ones that are incompatible with goals and plans, and only if more than one option remains do they try to identify the optimal one.

Other researchers have also pointed out that in real life, optimizing decision making is the exception rather than the rule. Features of the problem and the situation are often more important than the potential consequences of a decision. For instance, the behavior of managers in organizations is strongly determined by rules. The decisions of consumers are often determined by affects. In ethical conflicts, decisions are primarily determined by basic values, such as honesty and "do not harm," irrespective of the consequences.

Still another approach focuses on what is called naturalistic decision making, denoting the process by which people use their experience to make decisions in complex and dynamic environments, often under time pressure and involving high risks (e.g., pilot decisions on the flight deck). For such situations, in 1993, Klein proposed a recognition primed decision model that involves recognition of cue patterns that leads to retrieval of a response option. Serial evaluation of single options is considered typical, and the first option that matches the decision maker's goals and the situational constraints is chosen.

Models of organizational decision making describe decision making by both single and multiple actors in an organizational context. The normative rational (classical) model assumes that the organization follows a computational value-maximizing decision-making strategy. The model's usefulness is limited to a small set of situations with specific characteristics (e.g., goals can be described in quantitative terms); consequently, alternative models were developed based more on the actual behavior of decision makers in organizations. For instance, the (behavioral) organizational model focuses on the limited information-processing capacity of decision makers and postulates an outcome-satisfying strategy. In addition, the garbage can model conceives of organizations as "organized anarchies" (e.g., having inconsistent goals). Universities are considered the prototype of such organizations that do not follow one specific decision-making strategy. Outcomes result from the variable participation of their various members and groups in continually changing tasks. All of these models differ in their capacity to cope with different degrees of uncertainty and conflict among interests.

7. METHODS

Decision researchers use a variety of quantitative and qualitative methods to examine how people make decisions. Experimental studies in the lab usually ask participants to make judgments about values and importance as well as about uncertainties and risks, to make choices among given options, and to evaluate options, search for information about options, or distribute goods. Reaction times, eye movements, and other behavioral data (including verbal reports) are also used as measures of the information-processing activity of participants. Outside of the lab, the questionnaire is the dominant method, where either decision problems are presented by the researcher, and participants provide the required responses (e.g., students in a lecture) or decision situations arise and the researcher observes and records participants' behavior (e.g., pilots in a cockpit).

8. AREAS OF APPLICATION

The examples provided in this article for illustrative purposes have indicated some of the areas where decision making has been studied and where concepts and findings of decision research have been applied. More systematically, one may distinguish the following major domains: (a) behavioral economics and finance (e.g., investment and savings decisions); (b) business, administration, and management (e.g., employment, organizational, and product decisions; strategic planning with scenarios; communication of decisions; distributed decision making; cultural differences in decision making); (c) marketing and consumer behavior (e.g., product advertising, labeling, and pricing decisions; consumers' allocation of their scarce means over various commodities and services; consumers' decisions to acquire, consume, and dispose of products, services, and time); (d) health and medicine (e.g., addictive behaviors such as smoking, drugs, and alcohol; cancer screening campaigns; physicians' use of probabilistic information; diagnostic and treatment decisions; informed consent; shared decision making); (e) aviation, military command, and firefighting (e.g., pilot crew decisions and decision-aiding systems; battle commander decisions; general aviation safety decisions); (f) high-risk facilities (e.g., operator decisions in nuclear power and chemical plants; emergency and general safety-related organizational decisions); (g) environmental issues (e.g., individuals' benign attitudes and destructive behaviors;

governmental decisions and their implementation; communication between the public and the experts); (h) ethical and justice issues (e.g., organ transplantation decisions; business ethics; conflicts between self-interest and the other; dilemmas in risk communication); (i) advice, consulting, and counseling (e.g., decision advice in medical and financial problems; decision consulting for companies and agencies; genetic or abortion counseling for potential parents).

9. CONCLUSION

Approximately 50 years ago, the study of decision making started with lab experiments examining how people make choices among monetary gambles. Since then, the scope of studies has been extended immensely, reflecting the increasing importance of understanding and improving decision making in our society as well as the progress of methodology and model construction within the research community. However, the extension of the field has also eroded the idea of a general theory of decision making in favor of situation- and domain-specific models and empirical research on the conditions under which the various modes and strategies are being applied.

See Also the Following Articles

Leadership and Culture ■ Power, Authority, and Leadership

Further Reading

Beach, L. R. (1990). *Image theory: Decision making in personal and organizational contexts.* Chichester, UK: Wiley.

Connolly, T., Arkes, H. R., & Hammond, K. R. (Eds.) (2000). *Judgment and decision making: An interdisciplinary reader.* Cambridge, UK: Cambridge University Press.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*, 380–417.

Flin, R., Salas, E., Strub, M., & Martin, L. (1997). *Decision making under stress: Emerging themes and applications.* Brookfield, VT: Ashgate.

Hastie, R., & Dawes, R. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making.* Thousand Oaks, CA: Sage.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge, MA: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., & Tversky, A. (Eds.) (2000). *Choices, values, and frames.* New York: Cambridge University Press.

Koopman, P. L., Broekhuijsen, J. W., & Wiersma, A. F. M. (1998). Complex decision-making in organizations. In P. J. D. Drenth, H. Thierry, & C. J. de Wolff (Eds.), *Handbook of work and organizational psychology*, Vol. 4: *Organizational psychology* (2nd ed., pp. 357–386). Hove, UK: Psychology Press.

March, J. G. (1994). *A primer on decision making: How decisions happen.* New York: Free Press.

Messick, D. M., & Tenbrunsel, A. E. (Eds.) (1996). *Codes of conduct: Behavioral research into business ethics.* New York: Russell Sage.

Rasmussen, J. (1993). Deciding and doing: Decision making in natural context. In G. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 158–171). Norwood, NJ: Ablex.

Shapira, Z. (Ed.). (1997). *Organizational decision making.* New York: Cambridge University Press.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.

Decision Making in Sport

Gershon Tenenbaum

Florida State University, Tallahassee, Florida, USA

GLOSSARY

anticipation A mental operation that allows the athlete to predict upcoming events based on the elaboration between the current available visual–spatial information and knowledge stored in long-term memory.

coping strategies and self-regulation Active or passive emotion-focused or problem-focused methods used by athletes to appraise the environmental pressure that enable appropriate processing of information and decision making.

decision making A process by which an individual or a group of athletes select one course of action or actions from several alternative actions to be performed during precompetition or during a competition. It enables the athlete or the group to function under physically, emotionally, and mentally stressful conditions.

decision-making self-efficacy The belief that an athlete holds in his or her ability to make the right decision at the right time under competitive pressure.

perceived stress The perception of cognitive and somatic stimuli in the competitive environment that evoke feelings of anxiety and other emotions. It is linked to the degree with which one can adjust to the conditions to allow efficient decision-making processing.

visual–spatial attention The location in space in which the visual attention is focused, which enables a selection and elimination of environmental cues essential for decision making.

Jake plays in a central-midfield role on his soccer team. He plays at the very heart of the field and all moves, both offensive and defensive, bypass Jake. He has a multitude of decisions to make. Jake is young and prone to overarousal and constant errors. He pays equal attention to all the features of the soccer field and is limited in his capacity to process the most relevant information; thus, he often makes the wrong decisions. Coaches, parents, and teammates are upset with Jake; he is physically talented, but under time and event pressure he consistently chooses the wrong course of action. This article addresses the main issues related to Jack's limited cognitive capacity, which prevents him from making the right moves at the right time. The article addresses the required cognitive skills for proficient decision making and action execution. Several examples from different sports are provided to illustrate the research and theoretical issues.

1. INTRODUCTION

Decision making (DM) is a process by which an individual, a group, or an organization selects one preferred action from among two or more possible

actions within a specific situation. This article focuses mainly, but not conclusively, on the procedures involved in an individual athlete's DM processes. In sport, the DM process depends largely on the environmental and temporal conditions and rules under which a decision maker operates. These conditions vary with the nature of the sport. For example, in open-type sports such as basketball and soccer the environment is dynamic and consistently changing. The DM process in such sports requires the athlete to make a response selection (RS) based on past, present, and anticipated actions that may change through the course of time. Under such variable environmental conditions, the decision pertains to (i) which type of response/action to choose and (ii) at which time to execute the selected motor response. In contrast, closed skills, such as shooting and archery, which are performed within predetermined environmental constraints, require the decision maker to attend to his or her own proprioceptive sensory signals and decide when in the course of time to implement the decision.

In all sport environments, the DM process consists of three sequential phases: a preparation phase (i.e., visual search, selective attention, and anticipation); DM and RS through information processing (i.e., working long-term memory elaboration); and an action evaluation phase, in which DM and RS may be altered if new information is available at this stage. These stages are illustrated in Fig. 1. This information-processing model assumes that DM occurs under cognitive control. In cases in which the sport environment triggers a response under severe time constraints (i.e., less than 150 ms, which is the required reaction time under cognitive control), DM may occur under

self-organized rules. In such circumstances, decisions may come "naturally, without thinking." Generally, this self-organized DM process is more likely to occur among highly skilled and experienced athletes.

RS is typical to all sports, although the number of internal and environmental stimuli varies. The RS is more challenging when the number of stimuli increases, the temporal conditions shorten, and the athlete's affective state (i.e., anxiety) becomes less tolerable. The prevailing viewpoint in DM research in sport is that the process occurs within the context of both cognitive and affective systems, which operate simultaneously. The multifaceted perspective of DM in sport situations is presented in this article.

2. NONSPORT PERSPECTIVES ON DM

The traditional approach to DM was developed in the economics/statistics domain. The prescriptive approach viewed the person as an "outcome" seeker who pursues a desired goal. Accordingly, decision makers attempt to maximize a certain "expected value" and they process information in a manner that is targeted toward this ultimate end. The process is governed by comparing one's alternatives to an ideal solution that in many cases is unknown and speculative. One application of this approach accounted for gambling behaviors, where the unknown product necessitated a probabilistic estimation of utility. Despite their appeal, the prescriptive theoretical normative models were largely unsuccessful because they had limited ecological validity. Since human beings often make decisions that incorporate more considerations than just the expected value of their choice, these components needed to be identified and taken into account.

Cognitive-oriented researchers have examined the capabilities necessary for DM under certain environmental constraints. Also, cognitive capacity was believed to be a major constraint. Studies by DeGroot, Simon, Chase, and others in the early 1970s demonstrated that repeated exposure, experience, and skill level are associated with task-specific capabilities that enable the decision maker to use strategies and integrate information beyond the previously believed cognitive constraints. For example, expert chess players were able to recognize patterns in situations specific to chess in order to overcome limitations in working
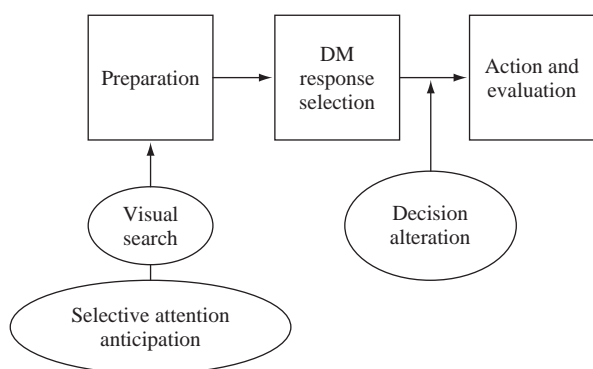

FIGURE 1 Stages of information processing and decision making in sport.

memory, although their overall cognitive capacities were not superior to those of nonexperts.

The naturalistic/descriptive approaches to DM assume that the DM process involves both rational and irrational processes and incorporates personal values, morals, motivation, and emotional level in the DM process. From a naturalistic standpoint, DM consists of defining the problem, seeking alternative solutions (some of which may not be available at a given moment), and then choosing an alternative. According to these approaches, the DM process can be compromised under various conditions, such as those imposing psychological pressure. Several theoretical models, such as image theory, explanation-based theory, recognition-primed DM, and cue retrieval of action, were introduced to account for the human DM process.

Contemporary models within the naturalistic/descriptive approach concentrate on how people make decisions rather than on how they should make decisions. Mechanisms in the form of heuristics are offered to account for the quality of decisions as well as the processes by which decisions are made. These models, which consider the real-world conditions under which decisions are made, have largely replaced the more mechanistic prescriptive models.

The recognition-primed model (RPD) developed by Klein assumed that past experience in similar situations governs the DM process. Klein suggested that DM consists of cue identification, situational goals, alternative action generations, and expectations for possible alterations, all of which can be enhanced with increased experience. The complexity of the situation generates more adjustments or alterations, when time permits. Mental representations, which are established through experience and repetitions, guide the DM process. RPD was found to be valid in chess playing, and this approach is closely related to current concepts and models in sport-related DM processes.

The naturalistic method is a knowledge-driven discourse (i.e., domain-specific declarative and procedural information processes) that according to Lipshitz has not been given its deserved attention. Recent studies emphasize the role that domain- and content-specific features play in utilizing decision strategies under varying environmental conditions. In sport, however, knowledge base in the form of expertise and deliberate practice is well documented, thus advancing the naturalistic conceptualization of DM in simple and complex conditions.

3. DM IN SPORT: THE DYNAMICS AND MECHANISMS

DM in sport constitutes a set of adaptive behaviors that enable an athlete to operate effectively within the competitive environment under conditions that vary with regard to emotional, mental, and temporal demands. The efficiency of this capacity is influenced by the richness and the variety of perceptions, as well as one's ability to process information, at a given time. Sport-related tasks are unique in nature and, therefore, encoding relevant information for further processing requires specific perceptual and attentional resources and mechanisms. Depending on the situation at any given moment, a shift between serial and parallel processing of environmental stimuli is required to allow an efficient RS process. Under unconstrained time conditions, information is fed forward for further processing, responses in the form of neural codes are selected and maintained "on alert," and a preferred action is selected at a given time while others remain activated to some degree until decayed. For example, a golfer may consider the weather and the terrain when selecting a particular club while awaiting his next shot. In contrast, under restricted time constraints the DM process operates under direct perception-action rules, which depend on knowledge structure and schemas stored in long-term memory (LTM). A "fast break" in basketball is a common example of this type of condition.

DM in sport can be viewed as a sequence of decisional processes rather than one RS. This sequence starts at an early stage when the athlete observes his or her opponents on a video film or uses his or her visualization skills. While watching the rival athlete's or athletes' style, weaknesses, strengths, and so on, the athlete plans potential actions, which are stored in memory in the form of representations that can be retrieved easily when required. The more the player is aware of the context, the more he or she can anticipate and alter his or her opponent's action even before stepping into the playing court. While competing, however, the first decision is related to the selection of relevant (and elimination of irrelevant) cues in the environment through the use of different visual strategies. A sequence of events in the performing environment is similar to a "mental chronometry" introduced by Posner, Nissen, and Odgen in which warning signals draw attention until an imperative signal triggers a response. During the preparation period, the athlete attends to the visual field and anticipates upcoming

events through the elaboration between working and long-term memory systems. The decisions at the anticipatory stage pertain to what actions the opponent intends to perform and with what probability (e.g., a tennis player anticipates the location of her opponent's serve). The next decision in the sequence regards what response to select from several alternative responses and when to execute it (where and when to move on the court). This decision has both quality and timing components. Last in the sequence is the decision about response alteration when conditions require or impose a change. This decision relies on a quick modification process, which controls the system's adaptation level to the changing and dynamic environment (how the player might recover if the serve goes in an unanticipated direction). The success rate of DM depends to a large extent on different mechanisms operating simultaneously and in mutual interaction. To clarify these processes, the main mechanisms affecting DM in sport are introduced next.

3.1. Visual–Spatial Attention

An athlete uses visual–spatial attention to select and discriminate among cues in the playing environment. In many sport settings, a vast array of stimuli, some relevant and some irrelevant, are available to the athlete, and a major aim here is to minimize environmental uncertainty and to enable smooth detection, recognition, recall, and selection of relevant stimuli for higher level processing necessary for RS. The visual system has two main characteristics: fixation, the target toward which visual attention is directed, and duration, the time lag between two shifts in fixation. Scientific evidence indicates that among athletes who play fast ball games, with increases in experience and skill level, the number of eye fixations decreases, whereas fixation duration increases. In other words, more experienced players tend to have fewer shifts in visual–spatial attention.

Visual scanning operates via two possible strategies: the target control strategy, which consists of scanning individual targets in a sequence until one is located that is compatible with a representation stored in LTM, and the context control strategy, which is initiated by memory representations not sensitive to individual objects but rather to a pattern in the visual display. Scanning the environmental stimuli under context control is of greater advantage in open and dynamic settings because it reduces the information-processing load, increases visual attention efficiency,

and simplifies the elaboration between working and long-term memory.

Novice players for whom the environment is not fully familiar adopt a target control strategy, whereas with increased experience and skill level a context control strategy becomes more prominent. With the shift from a target to context visual control strategy comes a reduction in eye fixations and an increase in the duration of fixations. In fast ball games, eye fixation location depends on the visual scan sequence. One first fixates on the large features in the display, such as the entire court in racket sports, then moves to the racket (tennis, badminton, and squash) during the later sequence before a RS is made. Such a synthetic strategy enables the player to integrate information from the core sources of reference at once instead of relying on serially analytic chronological order scan, which is time-consuming and inefficient.

Cueing and priming are also key features in the sport realm that are integral in visual attention. When an athlete anticipates that a move will occur with high certainty, his or her response is faster as a consequence of this "priming a response." When the environment causes greater uncertainty, response time is slower. While visually scanning the stimuli display, late cuing experiments in sport have shown faster reaction times to cue-primed locations than to unusual and unanticipated ones. However, sport-specific studies have shown that the cueing effect depends on the time interval between the warning and imperative stimuli. Reaction times increased with the time delay between the warning and imperative stimuli. However, in dynamic and fast-changing environments, perceptual-motor integration works differently from that reported for naive subjects in laboratory situations. As Cave and Bichot claim, the visual system is powerful in detecting and selecting objects in parallel to guide a response, and an object's location in the display is another criteria for DM. However, once extensively practiced, the probabilities assigned to upcoming events become the knowledge base of the athletes in the form of neural codes, from which they anticipate that events will occur. They thus become better equipped to prepare an appropriate action regardless of the time and space constraints of the environment. An example of this process is a batter in baseball who, through experience in detecting specific cues, learns which particular pitch (e.g., curve ball or fast ball) is most likely to be thrown in a certain situation.

An extensive knowledge base enables the athlete to minimize the costs associated with shifting attention

from narrow to large areas. When distracters in the space are cued, they interfere with the identification of a designated target and lead to an increase in reaction time. It is argued that when skilled athletes are exposed to several sources of information at one time, each carries with it a "probability value" for upcoming moves. When an anticipated action carries with it a high probability value, it enables the athlete to process the information efficiently and effortlessly and to subsequently make an adequate decision at the time. In contrast, an athlete who lacks such a capability pays the costs in the form of increased time for processing as well as less appropriate responses.

Finally, despite early beliefs that attention cannot be allocated or split among various sources simultaneously, recent research showed that attention could be divided when two targets were discriminated from each other by a given property or when all stimuli were suddenly presented. In open and dynamic environments such as ball games, anticipatory processes simplify the selection process by allocating visual attention to the more expected location and, at the same time, by using a context control strategy, which leaves other attention channels on alert for possible unexpected imperative stimuli to arrive. In such a manner, the skilled player secures a response with higher probability for success and at the same time leaves room for modification and refinement if circumstances require. These mechanisms help to explain how, for example, an elite basketball player can prepare to shoot a layup and, at the last moment, notice a defender's position, alter his decision, and pass the ball to a teammate.

3.2. Priming, Attentional Flexibility, and Dimensions of Attention

In the typical sport environment, cues are assigned a probability value in terms of response priming. That is, based on the visual information received, one response is given a higher probability of selection than the others. Responses are primed when their associated stimuli are detected. The primed response provides a "benefit" in terms of the reduced time required to process and select it, whereas the "cost" is that it then takes longer to process and prepare an alternative response for unexpected stimuli. The cost–benefit ratio determines the flexibility of the attention process.

Research using the expert–novice paradigm in sport has shown that the higher the skill level of the athlete, the more one is able to optimize this effect (i.e., to decrease the cost and increase the benefit of the stimuli-associated information processing). Precueing has a positive effect on the RS process because priming increases an athlete's awareness of upcoming events as well as response readiness. Scientific evidence also suggests that athletes can use automatic and voluntary signal detection modes and can shift between the two when required. With increasing expertise, the athlete acquires both a "smoother" shift and more pronounced flexibility process. In sport, the type and intensity of priming are essential for RS. In competitive situations, athletes choose the priming options that have the highest utility value; that is, they select those that ensure the best chances of triggering the appropriate responses at the proper times.

Because different sports require different attentional processes, no single strategy is ideal for all situations. Some sports (e.g., archery and golf) consist largely of stationary positions, requiring the athlete to focus attention on both internal and external signals. Sports that are more open and dynamic (e.g., basketball and football) require a wider and more flexible attentional style. Moreover, some sports include conditions that require a wide, external attention (e.g., a playmaker in basketball) but others that require a narrow internal attention focus (e.g., shooting a free throw). In these sports, attentional flexibility, which is the ability to alter one's focus, is a major asset. Neidefer's model of attentional style accounts for the requirements and shifts of attention needed in various sport contexts. Accordingly, he recognizes two dimensions of attention: (i) width, from wide to narrow, and (ii) direction, from internal to external. Athletes must acquire the skill of adopting the appropriate attentional width and direction that enables them to make effective decisions. Failure to shift attention increases the probability of missing the relevant information needed for optimal DM.

3.3. Anticipatory Mechanisms

The attentional system, which detects the most salient environmental cues, enables the athlete to anticipate events in advance. Repeated responses to game maneuvers enhance the athlete's perceptual anticipatory capability, which allows for a quick elaboration with the stored program representations in LTM. Fast access to one's knowledge base becomes possible with only partial information. Because practice in the form of task-specific skill develops ready-access scenarios, these preexisting schemes are available for access even when a limited amount of the information contained in

the game situations "yet to be" is captured by the visual system. This anticipatory capability becomes automated and effortless with the development of expertise. For example, an elite volleyball player may notice the movement of just one of six players on the opposing team. However, since she has a complex set of stored schema developed through experience, she is able to predict or anticipate the movements of the remaining players.

Scientific studies of fast ball games have demonstrated the advantage of expert athletes over their novice counterparts in anticipatory decisions. Advanced anticipatory capability reduced the time required for a response, particularly when temporal conditions were extremely short. Anticipatory capability stems from both knowledge structure and advanced visual attention strategies. During the initial stages of scanning the environment for relevant stimuli, anticipatory decisions alter depending on the events taking place at the time. In expert athletes, visual attention is directed to pattern recognition through enduring eye fixations, and fast eye movements occur only under certain environmental constraints. At this stage, attention is simultaneously diverted to several sources of information, including the opponent's moves, teammates' locations, and time left on the play clock. Lack of experience and target types of visual scan result in impaired anticipatory capability. At the final stages of the anticipatory process, the novice athlete is engaged in decisions about what and where moves and actions will be performed, whereas the expert makes decisions about what action to perform in response to the opponent's moves and/or situational demands. A novice wrestler, for example, decides where his opponent might step, whereas the expert selects the move he will use to score a takedown because he "knows" where the opponent's leg will be. Finally, RS appropriateness depends not only on the cognitive process but also on the proficiency of the motor system and the confidence one has in his or her ability to carry out the response. In the example cited previously, the wrestler must possess sufficient strength, quickness, and confidence for the decision to be effective.

3.4. Memory Representations, Information Delivery, and Knowledge Structure

When an athlete's knowledge base is rich, environmental distracters have limited impact on the DM process. A limited knowledge base considers these distracters as relevant information, resulting in slowed DM via inhibition of the athlete's information processing. The seminal studies of DeGroot and Chase and Simon on chess players in the 1960s and 1970s, and follow-up studies on a variety of sports since then, demonstrated how information is encoded, processed, and retrieved as a function of skill level and duration of visual exposure. It was assumed in these studies that recall accuracy following a short exposure to structured and unstructured situations would be a reliable indicator of both the visual attention strategy and the knowledge structure of the participants. Differences were consistent in that more skilled athletes could more accurately recall structured/logical/meaningful positions presented in a relatively short time than could unskilled athletes. The "chunking hypothesis" emerged as an explanation for this advantage in memory representation. This hypothesis rests on the premise that experience in the form of knowledge base and structure enables the player both to capture larger chunks of information in the form of logical patterns and to use these patterns as retrieval paths in a later stage. The chunking hypothesis was used to account for many research results, although recent scientific evidence from dancers and musicians has demonstrated significant memory superiority of skilled performers not only in meaningful information structures but also for memorizing random peaces of information. Thus, the chunking hypothesis has been called into question. In 1995, Ericsson and Kintsch proposed the existence of long-term working memory (LTWM), which may account for unusual memory capability in experts that is task-specific and cannot be generalized to other domains.

Short-term memory appears to have limited influence on experts' memory capabilities, and working memory was considered a constraining factor because only small amounts of information could be temporarily stored while other information was processed. The concept of LTWM overcomes the notion that memory storage capacity is limited and decays immediately after stimulus onset. The constraint of a limited capacity working memory system seems to be irrelevant in expert memory performance, wherein concurrent activity seems to present limited interference. The new LTWM conceptualizes that when one becomes an expert in a given domain, such as sport, he or she can utilize LTM as a means for expanding short-term working memory.

LTWM develops through the creation of a domain-specific retrieval structure that is used to enhance storage and to maintain information in a more accessible

and less interference-prone state. A specific knowledge base is developed through practice, which enables task-specific memory traces to be rapidly and efficiently retrieved. The advantage of LTWM is that it relies on fast encoding and storing processes in a ready retrieved form from LTM. Domain-specific stored information in LTM is used as a retrieval cue that allows one to "unpack" the requested retrieval structure. Although decisions in sport are required, these same memory traces are used to retrieve a response and keep another response alternative on alert for possible and immediate alteration, refinement, or modification. When the stored information is richer and readily accessible, the DM process is smoother and of higher quality.

For a RS process to be efficient, task-specific declarative and procedural knowledge should be structured in a form that allows for easy and fast access. French and McPherson explored the network of conceptual knowledge required to make and execute tactical decisions in the dynamic and changing environments of ball games. The knowledge base consists of memory representations for action profiles, game situation prototypes, competitive scripts, and sport-specific strategies. The knowledge base is structured hierarchically and consists of macro- and microlevel routes. The microlevel is managed by the working memory and enables the athlete to attend to ongoing events in the sport environment. Advanced environmental visual scanning and anticipatory capabilities enhance the flow of information for further processing. The macro knowledge level refers to action plan profiles and specific event profiles, which enable the athlete to control and regulate the sensory systems and efficiently select and execute a response.

3.5. Decision Alteration

Several studies incorporating overt (e.g., reaction time, choice reaction time, and accuracy) and covert behaviors (e.g., evoked potentials) simultaneously have indicated that correct and incorrect decisions can be detected in brain activity and described in neural codes. Although not examined in actual competitive situations, these results can be extrapolated into the sport realm. It is suggested that with practice and experience, the athlete incorporates a decision strategy so that he or she can successfully replace one response by another response when the conditions necessitate such an endeavor (i.e., decision alteration). The psychophysiological research clearly indicates that beyond a given neural activation threshold, response alteration is almost impossible. Nevertheless, to reduce the probability of

error, the expert athlete holds alternative decisions "on alert" and activates them when an alternative decision is required. Thus, expert athletes, unlike their nonelite counterparts, activate alternative pathways that allow an "effortless" alteration process, which occurs mainly in complex and time-limited conditions.

An explication of decision alteration among experts was presented in Shallice's "selection process" theory. According to this explanation, two attentional mechanisms control DM and action: contention scheduling (CS) and supervisory attentional system (SAS). The CS is automatic in nature and selects responses for action. The multiple schematic responses are activated and are in mutually inhibitory competition for selection. The final selection of a response is made when one of the alternative decisions is activated beyond a given threshold. The SAS mechanism can access the neural representations in the form of an intention for execution. When the selected response by the CS is proved to be inadequate during the time sequence, SAS modifies, stops, or, if not too late, alters the RS. Research shows conclusively that despite the advanced anticipatory capabilities of the expert athlete, RS is taking place only in the last stage of the process when uncertainty is minimized so that decision alteration can be made without tremendous processing costs. When an alternative response is ready and required, the alteration takes place.

4. DM IN TEAMS

One key advantage that a team has over an individual is a greater supply of human resources available for DM. For example, teams have an advantage over individuals on some measures of information processing (e.g., storage) owing to the extra information processing capabilities provided by multiple team members. Accordingly, in teams, the cognitive labor associated with DM tasks can be distributed over team members. Thus, the team can be considered as a cognitive system in itself. However, the cognitive properties of a team are more than just the sum of the cognitive properties of the team's constituent members; information processing is affected by the way the team communicates and is coordinated and organized. Consequently, DM efficacy increases to the extent that a team invests effort into establishing strategies for efficient communication, coordination, and organization.

Evidence supports the notion that the more a DM task requires intramember information exchange between team members, the more important are effective

communication strategies for performance. However, studies have revealed a tendency for teams that are under time pressure or that are newly formed to discuss only information shared by all team members and not information unique to individual team members. Teams may be subject to DM biases under these conditions. Similarly, studies on crew communication during ship navigation and subsequent use of computer simulations of communication have shown how confirmation bias, a phenomena known to affect DM in individuals, can be accentuated in teams. In 1991, Hutchins showed that even while holding the cognitive properties of individuals constant, groups may display quite different cognitive properties depending on how communication is organized within the group and over time. In summary, DM benefits from the extra resources a team can provide only to the extent that teams factors such as coordination, organization, and communication are taken into account.

5. AFFECTIVE STATES AND DM

DM in practice or competitive situations often occurs in the presence of psychological pressure and under conditions that elicit physiological arousal. Therefore, the cognitive process that leads to RS depends on the (i) perceived stress experienced at the time, (ii) coping strategies and self-regulations used to cope with the emotional and physical stress, and (iii) confidence the athlete has in successfully anticipating and selecting a response.

5.1. Perceived Stress and DM

DM is strongly affected by the perceptions of stress (i.e., anxiety) the athlete experiences during competition. Both somatic and cognitive anxieties affect the attention processes, which are vital for the DM to be effective. Under an excessively low-anxiety state, attention is directed internally, which may result in the athlete ignoring or avoiding crucial cues that, under an ideal state, would trigger a response. Under a high-anxiety state, attention narrows to a level at which important environmental cues are missed. The detrimental effects of an elevated anxiety level are most prominent in complex, open, and dynamic situations in which a plethora of environmental stimuli are required for effective DM. In closed and explosive-type sports, such as Olympic weight lifting, high levels

of perceived anxiety are less harmful and, indeed, can even be beneficial.

In line with major theories of cue utilization and attention narrowing under conditions of high perceived anxiety, various studies have shown that DM under low anxiety relies on processing relevant and irrelevant cues alike. Under such conditions, the most important cue is not always primed to trigger a response. Thus, selectivity under such a condition is unfocused and unlimited. Under high-arousal conditions, the opposite is true. That is, an athlete's attention is narrowed, and DM occurs with too few cues. Under moderate levels of anxiety, attention processes are optimal in that only the most relevant cues are primed for processing and DM. Highly anxious athletes more frequently conduct "false alarm" DM errors, (i.e., make unnecessary decisions), whereas athletes with low arousal levels conduct both false alarms and correct RS. Moderate anxiety levels also limit the problems that many athletes have concerning sustained alertness (i.e., vigilance). Optimal anxiety levels are associated with "alertness preservation" (i.e., high concentration) over a given period of time. Thus, when mental alertness is optimal, it enables one to process the environmental stimuli so that no major stimulus is missed.

A concern arises when "choking" under pressure is evident even in the expert athlete. Two theories account for DM and performance decrements under psychological pressure conditions. Distraction theory claims that once under pressure, shifting attention to task-irrelevant cues results in a move from automatic to conscious attention control, resulting in poor DM. Alternatively, self-focus theory or explicit monitoring theory proposes that pressure causes the performer to pay attention to and control the automaticity of the skill. When a skill is learned and is best executed under automatic control, a shift to intentional control negatively affects DM and motor response processing, resulting in choking. In short, the athlete "thinks too much." Increased arousal levels result in a narrowing and self-focused attention that shifts away from the relevant cues and ultimately affects the DM process. Furthermore, choking seems to be most evident in tasks that require procedural knowledge and less so in tasks that primarily consists of an explicit knowledge base.

A recent development with regard to the emotions–performance linkage indicates that RS depends on a larger range of emotions than merely cognitive and somatic anxiety. It is the individual's perceptions of the intensity, pleasantness, and functionality of emotions that determine the specific resources recruited,

which generate energy required for optimal DM. When such an emotional state is not experienced, DM, and subsequently performance, may decline. Thus, each individual has a unique emotional-related performance zone within which the probability of making sound decisions and implementing them is highest. In other words, optimal, moderate, and distracted decisions are more probable under different emotion zones for each athlete. These performance zones may overlap to some degree and may change with experience and skill level.

5.2. Coping Strategies and Self-Regulation

When emotions are elevated, coping strategies are used automatically as part of the process of protecting the self. Thus, different emotions produce different hormonal reactions, which trigger certain responses to cope efficiently with current conditions.

The methods athletes use to regulate themselves through the use, misuse, or avoidance of coping strategies determine how emotions are appraised and, by extension, the effectiveness and efficiency of the DM process. Repeated exposure to similar physically and emotionally stressful situations enables an athlete to better tolerate and monitor processes of attention and DM under pressure. Successful self-regulation can be attained and has been proven through the measurement of electroencephalographs, heart rate/pulse, and other physiological responses of the autonomic nervous system. Sustained alertness has been enhanced by avoidance of sleep deprivation; supplementation with nutritional ingredients; elevated physical conditioning; task-related strategies, such as breathing control, relaxation, imagery, and stress inoculation techniques; and conscious control over one's actions.

Coping has two basic forms: emotion-focused and problem-focused. Both strategies are aimed at changing the individual's appraisal of the current situation. An appraisal change results in a new meaning given to the situation, and attentional processes change as a result. Therefore, strategies such as avoidance, denial, and distancing imply one's intention to disengage from interfering stimuli to promote efficient use of visual attention, information processing, and other components necessary for effective DM. In other cases, when emotions are extremely elevated, such strategies may detract from information processing efficiency. Instead, approach strategies may be more helpful through the use of mental techniques such as relaxation, imagery visualization, and the setting of process and outcome goals. Research has consistently shown that the approach coping strategies are more helpful in monitoring the appraisal process under high perceived pressure, enabling the cognitive system to make decisions more efficiently. The more complex and dynamic the situation, the more beneficial their effect.

5.3. Self-Efficacy and DM

The social cognitive theory, particularly Bandura's self-efficacy theory, has important implications in sport DM. The beliefs athletes hold in their ability to make the right decision at the right time can be viewed as a coping strategy for handling decisions under competitive conditions. Self-efficacy in the context of sport DM has two dimensions: (i) the confidence one has in successfully anticipating and selecting the right response and (ii) the efficacy one has in executing the selected response. The limited research in this area leaves much to speculation. Current research indicates that self-efficacy in anticipatory decisions differs among novice, intermediate, and expert athletes. In fast ball games, experts display moderate levels of self-efficacy in their anticipatory decisions at early stages of response preparation. After the racquetball contact, anticipatory self-efficacy increases sharply. In novice and intermediate players, anticipatory self-efficacy remains moderate throughout the full duration of the task. These findings appear to support the notion that highly skilled athletes hold several alternative anticipatory decisions on alert under extremely short visual exposure frames, although this alters with the course of time—a pattern not typical of less skilled athletes.

6. SUMMARY: A HOLISTIC VIEW OF DM IN SPORT

DM in sport is a process of information processing in natural environments that impose unique constraints on the decision makers. When temporal conditions allow, the athlete relies on his or her experience and practice-related knowledge base to locate cues in the environment that are used for anticipatory DM. With experience and skill development, athletes shift from serial to parallel processing and from target to context visual strategies. These processes enable an athlete to

identify patterns of essential stimuli, elaborate on them efficiently via stored motor programs, and retrieve a response while keeping other responses on alert for possible decision alteration when needed. This process is efficient when the appropriate attentional style is employed throughout variable environmental constraints. The process of DM is more efficient when the athlete perceives the pressure of the sport environment as facilitative and less efficient when debilitative pressure is perceived. Being in the optimal emotional performance zone and feeling a high sense of efficacy for anticipating decisions and selecting responses increase the probability of making the correct decision. To secure an efficient DM process, the successful athlete employs the appropriate methods and mechanisms that underlie a skillful performance.

See Also the Following Articles

Coping ■ Decision Making ■ Stress

Further Reading

Abernethy, B. (2001). Attention. In R. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (pp. 53–85). New York: Wiley.

Bloom, B. S. (1985). *Developing talent in young people*. New York: Ballantine.

Ericsson, K. A. (1996). *The road to excellence: The acquisition of expert performance in arts and sciences, sports and games*. Mahwah, NJ: Erlbaum.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305.

Fisher, A. C. (1984). Sport intelligence. In W. F. Straub, & J. M. Williams (Eds.), *Cognitive sport psychology* (pp. 42–50). Lansing, NY: Sport Science.

French, K. E., & McPherson, S. L. (1999). Adaptations in response selection processes used during sport competition with increasing age and expertise. *International Journal of Sport Psychology*, *30*, 173–193.

Moran, A. P. (1996). *The psychology of concentration in sport performers: A cognitive analysis*. East Sussex, UK: Psychology Press.

Ripoll, H. (Ed.) (1991). Information processing and decision making in sport. *International Journal of Sport Psychology*, *3/4*, 187–406.

Starkes, J. L. (2000). The road to expertise: Is practice the only determinant? *International Journal of Sport Psychology*, *31*, 431–451.

Straub, W. F., & Williams, J. M. (Eds.) (1984). *Cognitive sport psychology*. Lansing, NY: Sport Science.

Tenenbaum, G. (Ed.) (1999). The development of expertise in sport: Nature and nurture [Special issue]. *International Journal of Sport Psychology*, *2*, 113–301.

Tenenbaum, G. (2003). Expert athletes: An integrated approach to decision-making. In J. Starkes, & A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 192–218). Champaign, IL: Human Kinetics.

Tenenbaum, G., & Bar-Eli, M. (1993). Decision making in sport: A cognitive perspective. In R. N. Singer, M. Murphey, & L. K. Tennant (Eds.), *Handbook on research in sport psychology* (pp. 171–192). New York: McMillan.

Tenenbaum, G., & Bar-Eli, M. (1995). Personality and intellectual capabilities in sport psychology. In D. Sakulufske, & M. Zeidner (Eds.), *International handbook on personality and intelligence* (pp. 687–710). New York: Plenum.

Tenenbaum, G., Levy-Kolker, N., Bar-Eli, M., & Weinberg, R. (1994). Information recall among skilled and novice athletes: The role of display complexity, attentional resources, visual exposure duration, and expertise. *Journal of Sport Psychology*, *12*, 529–534.

Tenenbaum, G., Levi-Kolker, N., Sade, S. S., Lieberman, D., & Lidor, D. (1996). Anticipation and confidence of decisions related to skilled performance. *International Journal of Sport Psychology*, *27*, 293–307.

Tenenbaum, G., Sar-El, T., & Bar-Eli, M. (2000). Anticipation of ball locations in low and high skill performers: A developmental perspective. *Psychology of Sport and Exercise*, *1*, 117–128.

Tenenbaum, G., Tehan, G., Stewart, G., & Christensen, S. (1999). Recalling a floor routine: The effects of skill and age on memory for order. *Applied Cognitive Psychology*, *13*, 101–123.

Williams, A. M., Davids, K., & Williams, J. G. (1999). *Visual perception and action in sport*. Routledge, London.

Dementia in Older Adults

Marie E. Cowart

Florida State University, Tallahassee, Florida, USA

GLOSSARY

cognitive function Intellectual abilities such as the dimensions of attention, language, memory, problem solving, orientation, visual–spatial abilities, judgment, and psychomotor skills.

cognitive impairment A general term referring to alterations in cognitive functioning that occur in some elders and that affect behavior and impair the ability to function in usual activities such as activities of daily living (ADLs); such alterations may or may not be reversible, depending on the causative factors.

dementia The permanent and progressive deterioration of intellectual function caused by changes in the brain that results in impaired thinking and altered behavioral impulses.

memory disorders A general term referring to the wide variety of structural and functional changes in the aging brain as well as a range of symptoms of memory impairments.

normal changes in aging cognition The normative changes in cognitive functioning as assessed from a population-based research perspective; there is a wide individual variability, with some individuals showing no changes at all.

Parkinsonism A chronic progressive disease of the nervous system characterized by a slowly spreading tremor, muscle weakness, and (during the later stages) dementia.

Pick's disease A brain disorder involving atrophy of the cerebral cortex and resulting in asthenia (i.e., loss of strength), loss of speech, and progressive dementia.

senescence A general term referring to the process of growing old.

senile dementia (or Alzheimer's disease) A progressive disease characterized by neuritic plaques and neurofibrillary tangles that leads to cognitive decline; it is most common during older adulthood.

senile plaques Extracellular brain tissue changes described by Alzheimer in 1907 and found in the brains of persons with Alzheimer's disease and, to a lesser extent, in normal-aged brains.

senility A term referring to a generalized deterioration in mental or cognitive skills, leading to a decline in the ability to perform social or occupational activities.

vascular dementia The second most common dementia, caused by an interruption in cerebral blood flow due to stroke.

This article deals with the concept of senility using a broad frame of reference. Cognitive effectiveness and age-related cognitive decline or impairment have been the subject of much research, especially during the past decade or so. At this time, there is not a consensus regarding all ideas about screening for dementia and preventive treatment to delay progression of the

disease. During the coming decade, more research findings will help to answer questions about the most reliable screening methods and the efficacy and ethics of using them. Controlled longitudinal studies will help to determine the value of pharmacological and nonpharmacological interventions. This article explores the meaning of senility, what it is and what it is not, who is at risk, prevention, early and late recognition, comorbidities and misdiagnoses, interventions, and future prospects.

1. SENILITY: WHAT IT IS AND WHAT IT ISN'T

1.1. History

Senile dementia was mentioned in the writings of Hippocrates and Pythagoras during the Greco–Roman era. It was recognized by Bacon in 1290, by dePratis during the 1500s, and by Pinel during the 1800s when he reformed the treatment of the mentally ill. At the turn of the 20th century, Alzheimer defined clinical and physiological brain lesions of senile dementia, now referred to as Alzheimer's disease.

1.2. Definition of Normal and Impaired Cognitive Functioning

It is important to differentiate between normal changes in aging cognition and senility or dementia. In normal aging, cognitive processes are slower and less accurate for more effortful or novel tasks than they are during younger years. Therefore, the older person may take longer to absorb the communication message, retrieve a thought, and respond. Memory is a special problem in aging because the older person takes longer to store information so that it may be recalled. Some have likened the aging brain to an overworked computer. The processor of the older computer is slow and the hard drive is overloaded with stored data, thereby making retrieval a more complex and longer process. Sometimes the files get disorganized, particularly when the computer is tired and overused. The new computer has a fast processor and little data stored in it so that retrieval is rapid. In terms of memory retrieval and learning new information, the older person processes information more slowly than does the younger person. The older person's cognition may be likened to the old computer's processor.

Working memory becomes less efficient in the older person. Working memory may be thought of as the ability to retain or temporarily store some information while at the same time accessing or using other information to perform a task. A complex environment, such as a loud and busy room, may interfere with working memory performance in the elderly. Working memory is needed in problem solving.

In terms of problem solving, the older person has more experiences to draw on, so decision making may be more effective in the older person than in the younger and less experienced person. Some refer to this as judgment or wisdom.

In dementia, there is a breakdown in several dimensions of cognition. Judgment is impaired and memory function is disrupted in terms of storage retrieval and planning for future conversations or decisions. Communication breaks down because the person is unable to problem solve or think ahead about the flow of words to use. Confusion and frustration may result, leading to behavioral and emotional changes.

In 1994, the American Psychological Association defined dementia as having the following deficits: memory impairment, cognitive disturbances (e.g., aphasia, apraxia, agnosia), and disturbances in executive functioning (i.e., goal setting) that result in significant impairment in social or occupational functioning and represent a significant decline from a previous level of functioning. According to this widely accepted definition, senility or dementia becomes a disability when social or occupational functioning is impaired.

2. MYTHS AND STEREOTYPING

Unfortunately, "ageism" or labeling sometimes can cause an older person to be called senile or demented when perhaps he or she is only hard of hearing or has received a large volume of information rapidly so that there is not time to encode it for memory retention and recall. Such stereotypes are common and permeate our society.

Consider the example of an 85-year-old and a 35-year-old talking with a group about a party they had attended the previous night. Both met a number of new acquaintances, but each could recall only one or two new names. The older person was quickly labeled senile, whereas the 35-year-old was viewed as just not having paid attention.

3. EXTENT OF THE PROBLEM

There is a clear association between cognitive decline and aging. This association is confirmed with cross-national prevalence data. Because of longevity, an increased percentage of the population is seen as having cognitive decline compared with earlier generations when the life expectancy was shorter. In 1996, Sloog and colleagues reported that 1% of 65- to 70-year-olds are diagnosed with moderate to severe memory disorders, whereas that number increases to 21% in 85-year-olds. Danish studies published in 2001 reported that among centenarians, 21% show no dementia symptoms, whereas 40 to 63% have mild to severe dementia.

3.1. Positive Risk Factors

There is much discussion that keeping an active mind deters memory and other cognitive function decline, although these approaches have yet to be tested longitudinally with large samples. Prevalence studies do indicate that a higher educational level, high occupational status and workplace complexity for older ages, an intact marriage to a well-educated spouse, a stimulating environment, and engagement in cultural and educational activities are associated with higher intellectual functioning into old age.

3.2. Negative Risk Factors

A number of disease conditions increase the risk of cognitive decline in individuals. Conditions such as a previous head injury, Parkinsonism, and multiple small brain infarcts or hemorrhages are common conditions that are likely to progress to cognitive impairment.

Major risks for dementia or cognitive decline are chronic diseases. Longer life expectancies are accompanied by increased prevalence of chronic conditions, resulting in more dementia in old age today than was the case during earlier decades when life spans were shorter.

Withdrawal from intellectual engagement, such as retirement from an intellectually challenging and complex job, can be a negative risk factor. These and other associations with cognitive functioning provide positive arguments for continued intellectual activity into old age as a deterrent or delay strategy for dementia.

There is some disagreement in the literature as to whether a busier and more complex life can lead to

fatigue that interferes with cognition or, if kept in balance, can provide intellectual stimulation. A lifestyle that allows for self-pacing provides more time for reflection and provides an opportunity to encode new information and retrieve old information without hurrying.

4. PREVENTIVE MEASURES

The research literature is clear in that there are no absolute preventive measures for senility or age-related dementia. If genetic predisposition occurs, preventive measures might only delay symptoms. However, studies have shown repeatedly that prevalence rates of dementia are lower in cases where certain associations have occurred. Some of these are discussed in this section.

4.1. Nutrition and Nutritional Supplements

Various studies have shown associations between the prevalence of Alzheimer's disease and the content of food intake. Vitamins that fall into the antioxidant category are considered to be beneficial in preventing dementia due to their ability to repair or replace molecules that are damaged by free radicals. The National Institute on Aging currently supports several large clinical trials that look at the effects of Vitamins A and C, selenium, betacarotene, folate, and multivitamins on cognitive decline and on the prevalence of Alzheimer's disease. The outcome of these large trials will provide additional empirical basis for evaluating the effect of vitamins and food supplements on aging cognition.

Reduced intake of saturated fat and maintenance of a low cholesterol level are believed to lower the risk of the onset of dementia and Alzheimer's disease. Studies have shown some association of lowered prevalence of cognitive decline among elders who eat a diet that is high in fish and other unsaturated fats. Fish oils taken in large amounts are reported to have a positive effect on the retention of cognitive function. These and other food substances have not been well tested with controlled and longitudinal studies on large samples.

4.2. Pharmacological Effects

Repeated studies have also shown some positive association between cognitive functioning and long-term use of nonsteroidal anti-inflammatory drugs such as

the regime that an individual with arthritis might take. People on prednisone, a steroidal drug, have not shown protection against dementia.

There have been a number of studies on estrogen showing mixed results regarding its effect on protecting cognitive functioning in women. The pattern seems to be that the larger trials, such as the Women's Health Study, show that there is no protective effect of hormone replacement therapy in older women on the prevalence of dementia and Alzheimer's disease.

4.3. Physical and Mental Activity

The concept of neuroplasticity applies to the ability of the adult brain to remodel pathways as occurs in the rehabilitation process in persons with stroke or brain injury. Magnetic resonance imaging has helped researchers to visualize brain mapping and to observe the reconfiguration of active pathways. Although confirming research trials have not been conducted to show the brain's ability to remap cognitive and memory function, the concept of "use it or lose it" has gained strength during recent years. Studies showing associations of cognitive function with higher education levels, as well as greater intellectual and social activity, have provided incentives that encourage people to engage in cognitive activities throughout their lifetimes so as to prevent age-related dementia. Although more longitudinal studies are needed, recommendations to engage in lifelong learning, even into old age, are widespread. However, the lasting effects of such behavior are not known.

5. EARLY AND LATE RECOGNITION

Ethicists have discussed the efficacy of screening for dementia. Who decides to screen? What is done with the information? Who decides to share the results of dementia testing? What are the implications for labeling a person with the Alzheimer's disease diagnosis? Who has the right to refuse screening? Does screening make a person disability eligible, or do the results of screening disqualify a person for the workplace or even for health insurance? Before screening for dementia, clinicians should ask themselves these and other questions and discuss the implications of screening with their colleagues and patients.

5.1. Early Recognition and Intervention

In looking for signs of the onset of dementia, it is important to differentiate between normal changes in memory that occur with aging and changes that may be associated with dementia such as Alzheimer's disease. In normal aging, the older person takes longer to learn new information, and the encoding process requires more time for new material to be retained, than is the case in younger persons. Therefore, an apparent inability to learn or to remember what was just heard is not necessarily a sign of early dementia; rather, it may be a sign of slowed cognition that occurs in all aged persons.

Changes with dementia, on the other hand, emerge with difficulty in learning new information. The person often gives up trying to learn. Difficulty in learning subsequently affects the ability to recall, that is, memory. This applies to verbal information as well as spatial information such as remembering directions.

5.2. Clinical Picture

The person with early age-related dementia may be unable to initiate new ideas into conversation, appears to be repetitive in speech, engages mainly in familiar routines to minimize stress, and shows behavior changes such as impatience, frustration, and withdrawal or depression. The person may encounter difficulty in concentrating as well. Complex skills, such as language planning and problem solving, may be shunned if the person with early impairment is aware of his or her own difficulty in cognitive functioning. Later stages show more memory impairment and language skill deterioration. The advanced stage of dementia will show impairment of all mental functions, including changes in personality and emotions exhibited by changes in behavior.

5.3. Screening and Diagnostic Tools

It is important that an assessment of an individual's cognitive functioning not be based on one instrument or tool. A multifaceted assessment of cognition will require an interdisciplinary team approach because special training is necessary to administer these and other approaches to evaluation. Where teams are absent, the use of a variety of assessment tools is

important. A basic assessment regime might include the following procedures:

- History and physical examination
- Neuroimaging and electroencephalography (EEG)
- Blood and spinal fluid laboratory tests
- Social history from a close family member
- Pharmacological and food supplement history
- Evaluation of activities of daily living (ADLs) and instrumental activities of daily living (IADLs)
- One or more of the following cognition data collection instruments:

1. *Mini-Mental Status Examination.* The Mini-Mental Status Examination (MMSE) is a common screening test that is used because of its reliability and validity in evaluating dementia, although it may show some bias for education and socioeconomic status. The 30-item instrument covers orientation, registration, attention, calculation, short-term recall, language, and construction. The test is not sensitive to early dementia.

2. *Clock Drawing Test.* In the Clock Drawing Test, the patient is provided an examiner-drawn circle and asked to put numbers in it and then to draw the clock hands at a given time (e.g., 11:10, 2:45). The properly administered test requires abstraction. Although it is sensitive, it should not be used alone as the sole screen or measure of cognition.

3. *Depression scales.* Although a number of depression scales are available, two seem to have common applicability for the elderly. The Geriatric Depression Scale (GDS) has several versions (5–15 items per version) and is a self-report of how the individual has felt during the past week. The Beck Depression Inventory (BDI) is a 21-item self-report scale that includes some somatic symptoms that may also be found in other chronic conditions. A depression assessment is important when determining dementia in older persons due to the frequency with which depression occurs in the elderly and concern that signs of depression may mimic or cloud those of dementia.

4. *Activities of daily living.* An important part of the dementia assessment is an evaluation of the individual's ability to carry out routine ADLs such as bathing or showering, dressing, eating, transferring, and toileting. Each item is rated as minimal assistance, moderate assistance, maximal assistance, or total assistance. This information may be gained from direct observation, interview of the caregiver or close relative, or self-report. A measure of ADLs can be used to rate the severity of the psychomotor or behavioral accompaniments of dementia.

5. *Instrumental activities of daily living.* The IADL tool provides higher level functions than does the ADL tool. IADLs include food preparation, home maintenance, laundry, financial management, communication (including use of a telephone), use of transportation, and medication administration. The IADL tool has been criticized for having a gender bias due to the food preparation and household items.

6. MISDIAGNOSES AND COMORBIDITIES

Careful screening and diagnosis is important in determining the presence of dementia, particularly during the early stages when signs and symptoms might not be obvious. Vision or hearing problems can interfere with verbal or spatial input and can impair the individual's ability to communicate accurately, thereby obscuring some cognitive functions such as the date and time. Sorting out such conditions is facilitated by multidisciplinary evaluation and repeated assessments over time to ensure an accurate diagnosis.

Frequently with the onset of multiple chronic conditions and diseases, it is difficult to determine whether comorbidities precede or follow the onset of dementia. This is particularly true in the case of senility or dementia. There are some frequently observed conditions that are strongly associated with dementia or are known to predispose to cognitive dysfunction. A number of common ones are mentioned in this section.

6.1. Head Injury

Depending on the location and severity of a head injury, or in repeated head injuries such as those that boxers may experience, dementia can occur.

6.2. Stroke, Atherosclerosis of Cerebral Arteries, High Cholesterol, and Hypertension

Vascular dementia is an outcome of some strokes or advanced atherosclerosis of cerebral arteries causing cerebral emboli. Persons with advanced diabetes can experience vascular dementia.

6.3. Oxygen Depletion of Cardiovascular or Respiratory Origin

Decreased cerebral oxygenation can lower brain metabolism and, over time, can lead to cognitive impairment.

6.4. Parkinsonism

During the advanced stages of Parkinsonism, this neurodegenerative disease can exhibit cognitive impairment, particularly in working memory.

6.5. Chronic Mental Illness

Some long-term mental health problems may show a slow emergence of dementia that may be due to the mental condition or to the effects of long-term therapy.

6.6. Drugs or Other Toxic Substances

Long-term exposure to certain toxic substances can affect cognition. Short-term drug toxicity can cause the person to exhibit changes in cognitive function, particularly orientation, that can be reversible in many instances.

6.7. Other Conditions Such as Dehydration, Fatigue, and Relocation

A fair amount of research points to the effects on working memory and orientation when the older person experiences unusual physical or mental stressors such as those that may occur with dehydration, extreme fatigue, and sudden relocation of the living environment. Research has demonstrated that older people have times of peak functioning during the day, usually the morning hours. Careful assessment of the entire person is essential before applying the label of senility or dementia.

7. INTERVENTIONS AND CARING

There is a great deal of literature on interventions for individuals with moderate and advanced dementia. This article deals with some early interventions and concerns for the care of persons with cognitive impairment.

7.1. Dietary Supplements and Medications

Aricept (other trade names include Cognex and Exelon) is the most widely used drug to alleviate the symptoms of Alzheimer's disease. It does this by prolonging the effects of acetylcholine release events, thereby raising neurotransmission and increasing neurotransmission effectiveness. Its use is thought to delay the advancement of symptoms rather than to provide a cure, although side effects may interfere with use of the drug.

7.2. Memory Aids and Other Nonpharmacological Treatments

The use of memory cues, such as written notes, is thought to help prolong independence in people who have progressive age-related dementia. Some literature indicates that people with mild or even moderate dementia retain reading skills longer than they do oral skills. There is also evidence that the use of written language becomes less complex in terms of sentence structure and vocabulary as the dementia progresses.

If familiar routines are retained in persons with mild to moderate dementia, independent living can be extended. Conversely, learning new routines can trigger undesirable or stressful behavior. Community-based group programs are often available through a day care format and use a variety of activity approaches to enable the individual to retain a desirable quality of life and to provide respite for the family caregiver.

7.3. Management Concerns

As the dementia condition progresses from mild to moderate levels of impairment, a number of concerns that need to be addressed may arise. Because dementia emerges in very different ways for each individual, not everyone with moderate or advanced dementia will experience these problems. However, an awareness of possible concerns will enable the practitioner to aid the family caregiver in planning ahead to prevent unsafe or disturbing behaviors.

Driving is an activity that is closely identified with a person's sense of autonomy and independence. Yet as the individual's judgment and problem solving become impaired, it is not unusual for the person with

dementia to become lost or to make wrong decisions that can result in an auto accident. Determining when it is no longer safe for the individual to drive is a problem that family members face and that often requires their intervention.

Wandering is another concern for the ambulatory person with moderate or advanced dementia. Assistive devices, such as radio and identifier tags, can help to quickly locate the individual who has the potential for wandering. Failure to locate the person who has wandered may lead to dehydration, hunger, hypothermia, and even death.

When the person with moderate dementia is unable to carry out two or more ADLs, additional support for the family might need to be considered. Such support may be provided by adding additional assistance in the home. If the behavior is unmanageable, a time may come when the family decides to move the individual to a protective environment such as an assisted living facility for persons with dementia. Measures of family coping may assist in determining the amount of caregiving burden that the caring family members are experiencing and may serve as a useful tool in assisting the family with making decisions about care.

Repetitive forgetfulness can pose problems for caregiving. Forgetfulness can include leaving a stove on, leaving a cigarette lit, and not being able to locate familiar objects such as keys.

These are just a few of the concerns that can arise in providing care for the person with moderate dementia. The caregiver can gain many tips on how to deal with caring for the family member with dementia by joining a support group and sharing concerns with others who face similar experiences.

8. FUTURE PROSPECTS

Many call the decade of the 1990s as the "decade of the brain." Although much knowledge has been gained about brain function, there is much more to learn. Further advancements in maintaining cognitive effectiveness are certain to be made, although a cure for age-related dementia is not anticipated. Whether new approaches, such as brain stimulation from electrical or pharmaceutical products, are developed and are effective in lessening the progression of dementia is another area for empirical study.

Recent brain research, summarized in the September 2003 issue of *Scientific American,* indicates that there is evidence that brain pathways for motor control have

the capacity to change. This phenomenon is referred to as plasticity. Future research will demonstrate whether repatterning can occur in intellectual functioning as it has with physical activity.

Impending demographics in terms of increased life expectancy and population aging means that there will be increasing numbers of elders throughout developed and newly industrialized nations. Greater numbers of people living longer will mean that increases in the prevalence of degenerative diseases, including dementia, will be experienced throughout the world. If preventive approaches and proposed therapies are found to be effective, much will be saved in terms of human and economic resources, and the quality of life of large numbers of elders will benefit.

See Also the Following Articles

Aging and Competency ■ Aging, Cognition, and Medication Adherence ■ Cognitive and Behavioral Interventions for Persons with Dementia ■ Cognitivism ■ End of Life Issues ■ Psychotherapy in Older Adults

Further Reading

Blackman, L., Small, B. J., Wahlin, A., & Larson, M. (1999). Cognitive functioning in very old age. In F. I. M. Craik, & T. M. Salthouse (Eds.), *Handbook of aging and cognition* (Vol. 2, pp. 449–558). Mahwah, NJ: Lawrence Erlbaum.

Chen, P., Ratcliff, G., Belle, S. H., Cauley, J. A., DeKosky, S. T., & Ganguli, M. (2001). Patterns of cognitive decline in presymptomatic Alzheimer's disease: A prospective community study. *Archives of General Psychiatry, 58,* 853–858.

Elias, M. F., Beiser, A., Wolf, P. A., Au, R., White, R. F., & D'Agostino, R. B. (2000). The preclinical phase of Alzheimer's disease: A 22-year prospective study of the Framingham cohort. *Archives of Neurology, 57,* 808–813.

Fratiglioni, L., & Rocca, W. (2001). Epidemiology of dementia. In F. Boller, & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 6, pp. 193–215). Amsterdam: Elsevier.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 58,* 2034–2039.

Pope, S. K., Shue, V. M., & Beck, C. (2003). Will a healthy lifestyle help prevent Alzheimer's disease? *Annual Review of Public Health, 24,* 111–132.

Rabbitt, P. (1998). Aging of memory. *In Brocklehurst's textbook of geriatric medicine and gerontology* (5th ed., pp. 123–152). New York: Churchill Livingstone.

Smith, A. (1996). Memory. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 236–250). San Diego: Academic Press.

Depression in Late Life

Barry A. Edelstein, Andrea Shreve-Neiger, and Stephanie A. Scheck

West Virginia University, Morgantown, West Virginia, USA

GLOSSARY

age-related changes Normal physical and psychological changes from young adulthood through old age.

comorbidity The co-occurrence of two disorders or syndromes.

late-onset depression Depression with an onset at age 65 or older.

major depression Major depressive disorder as defined in the *Diagnostic and Statistical Manual of Mental Disorders*. It is characterized by depressed mood for at least 2 weeks and at least four additional symptoms of depression.

minor depression Minor depressive disorder as defined in the *Diagnostic and Statistical Manual of Mental Disorders*. It is characterized by depressed mood for at least 2 weeks but with few symptoms and less impairment than in major depressive disorder.

reminiscence therapy Involves recall of events in a person's life and is usually done in groups in which participants are encouraged to talk about past events, sometimes with such aids as photos, music, and personal objects.

somatic symptoms Fatigue, changes in appetite, changes in sexual activity, and changes in sleep duration.

subsyndromal symptoms Symptoms of depression that are not of sufficient frequency, duration, or intensity to meet the formal criteria for an affective disorder.

The 1999 Surgeon General's Report on Mental Health revealed that almost 20% of older adults experience mental disorders. Depression, although not the most prevalent mental disorder among older adults, is a very significant source of suffering. From a clinical perspective, it is a particularly complex and challenging phenomenon, due in part to age-related differences in the experience and presentation of depressive symptoms, risk factors, and the many potential causes of depressive symptoms (e.g., medications, comorbid diseases, and stroke).

1. EPIDEMIOLOGY OF LATE-LIFE DEPRESSION

Depression in older adults is often overestimated and underrecognized, and it is not the most prevalent psychiatric disorder among older adults, as many believe. Indeed, the prevalence of depression is lowest among adults aged 65 or older. Prevalence rates for major depressive disorder among community-dwelling older adults (65+ years) range from 1 to 4%. The North Carolina Epidemiologic Catchment Area Study estimated a prevalence of 0.8% for major depression, 2% for dysthymia, and 4% for minor depression. Incidence

estimates of depression in long-term care facilities have ranged from 12.4 to 14%. If one considers clinically significant depressive symptoms that do not meet diagnostic criteria (subsydromal symptoms), the rate of depression ranges from 8 to 15% for community-dwelling older adults, 25 to 33% for hospitalized older adults, 10 to 15% for nonpsychiatric outpatients, and 30 to 40% for long-term care residents. The Epidemiological Catchment Area Studies reported bipolar I and bipolar II prevalence rates of approximately 0.1% among older adults.

2. CONCEPTUAL AND DIAGNOSTIC ISSUES

There is increasing evidence that depression among older adults differs from that of younger adults along several dimensions (e.g., etiology, presentation, and phenomenology). In 1988, Alexopoulos *et al.* suggested that the etiology of depression among older adults appears to be somewhat related to the age of onset. Gatz and Fiske noted that among depressed older adults, approximately 50% are cases of late-onset depression. The etiology of late-onset depression may be quite different from that of depression occurring in earlier years. The experience and presentation of depression among many older adults appears to be different than that of younger adults. For example, depressed older adults are less likely to report dysphoria than younger adults with the same level of overall depression. Gatz and Hurqitz revealed that older adults are also more likely than younger individuals to endorse test items suggesting a lack of positive feelings, whereas younger adults are more likely to endorse test items suggesting depressed mood. Older adults are also less likely to report ideational symptoms (e.g., guilt) than younger adults and more likely to report somatic symptoms. These somatic symptoms, such as fatigue and changes in appetite and sexual activity, may be related to coexisting medical conditions or normal age-related changes. Reports of lack of interest, sleep disturbance, and lowered energy level, however, may be useful in differentiating depressed and nondepressed older adults.

In addition to the content of age-related symptoms of depression, there is evidence that symptom levels and patterns that would not currently meet *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (*DSM-IV*), criteria for major disorders are nevertheless problematic for older adults. Many older individuals present with somatic complaints and experience symptoms of depression that do not meet the criteria for depressive disorders. The effects of these subsyndromal conditions may be just as consequential as those of the disorders that do meet diagnostic criteria.

In summary, depression among older adults appears to be different in a number of ways from that encountered in younger adults. The etiology of depression among older adults appears to vary depending on the age of onset. Moreover, the nature and pattern of depressive symptoms of older adults appear to deviate from those one would expect based on depression criteria found in the *DSM-IV*.

3. RELATIONS AMONG DEPRESSION, PHYSICAL DISEASES, AND MEDICATIONS

One of the complicating factors faced by clinicians and depression researchers regarding the nature and etiology of depression among older adults is that older adults have a higher incidence of medical disorders than younger adults. Approximately 80% of older adults suffer from at least one chronic health problem. Consequently, it is relatively common for depressed older adults to have comorbid medical disorders; that is, they exhibit the co-occurrence of two disorders or syndromes.

Although the relations between depression and comorbid disorders can be rather complex, they are generally well established. Depression can precede, follow, or simply co-occur with unknown directional relations. Depression is a risk factor for physical disease (e.g., cardiovascular disease) and diminished cell-mediated immunity. Symptoms of depression can also arise directly from various physical diseases (e.g., hypothyroidism, stroke, Parkinson's disease, and pancreatic cancer) and from the psychological reaction to a disease (e.g., Alzheimer's disease). Depressive symptoms may also follow a disease process (e.g., myocardial infarction). Similarly, depression is very common among cancer patients, with rates varying from 25 to 50%; there is a greater prevalence among individuals experiencing higher levels of pain, physical disability, and illness. When considering the relations between physical disease and depression, it is important to understand that one need not meet criteria for depression for the symptoms of depression to affect physical health.

The complexity of the relations between depression and physical diseases is often amplified by the adverse effects of medications used to treat medical diseases (e.g., parkinsonism, pain, cancer, arthritis, and hypertension) because many of these medications can produce symptoms of depression. Older adults consume more medications than any other age group and are at higher risk of adverse drug reactions than any other age groups because of age-related changes in physiology and increased use of multiple medications.

4. ASSESSMENT OF DEPRESSION

The assessment of depression among older adults can be quite challenging, particularly in light of the issues discussed previously. Age-related differences in the presentation of depression as well as the potential for coexisting physical aliments, medication side effects, and natural age-related changes increase the complexity of diagnosing depression when assessing older adults.

Methods of assessment for depression include self-report, clinical interview, report by others, physical examination, and direct observation. Given the intricacy of identifying depression in older adults, there may be no single best method for assessing depression. Each method has its own relative strengths and weaknesses; however, combining methods during the assessment process can enable a clinician to obtain a more complete understanding of depressive symptomatology in older adults. Here, discussion is limited to the most commonly used self-report and clinician-rated instruments.

5. INSTRUMENTS FOR THE ASSESSMENT OF DEPRESSION

The most widely used instruments for assessing depression use the self-report method due to the relative ease of administration and brevity. Clinician rating scales, although often taking slightly longer to administer, offer the advantage of objective observation of the depressive symptoms, which can be particularly useful with cognitively impaired older adults.

5.1. Self-Report Instruments

5.1.1. Beck Depression Inventory

The Beck Depression Inventory (BDI) is a 21-item, multiple-choice inventory. Respondents are asked to rate each item based on four response choices according to the severity of the symptoms, ranging from the absence of a symptom to an intense level, during the past week. A 13-item version of the BDI is also available. Both the original and short forms have reasonable internal consistency for normal and depressed older adults and adequate test–retest reliability in older adult patient and nonpatient populations.

A second version of the BDI (BDI-II) was published in 1996 so that items would be consistent with the revised criteria for depressive disorders in the *DSM-IV*. The number of items on the BDI-II is the same as that for the original version. Edelstein *et al.* suggest that the BDI-II be used with caution with older adults given the lack of empirical support for its use with this population.

5.1.2. Center for Epidemiological Studies–Depression Scale

The Center for Epidemiological Studies-Depression Scale (CES-D) is a 20-item instrument with good psychometric properties. A 10-item version of the CES-D also has adequate psychometric properties and appears to be a good instrument for screening for depression in community-dwelling older adults.

5.1.3. Geriatric Depression Scale

The Geriatric Depression Scale (GDS) is a 30-item scale designed specifically for use with older adults. The scale uses a yes–no format and has well-established psychometric properties with community-dwelling older adults, outpatient and hospitalized older adults, and geriatric stroke patients. A 15-item form of the GDS yields scores that are strongly correlated with those of the long version. Overall, the GDS appears to be a useful instrument for assessing depression in older adults.

5.1.4. Zung Self-Rating Depression Scale

The Zung Self-Rating Depression Scale contains 20 items that are rated on a 4-point scale by the patient according to the amount of time he or she currently experiences the symptom. Psychometric properties with older adults are less than desirable.

5.2. Clinician Rating Scales

5.2.1. Hamilton Rating Scale for Depression

The Hamilton Rating Scale for Depression (HRSD) is used to quantify the severity of symptoms of depression and is one of the most widely used and accepted instruments for assessing depression. The standard version of the HRSD is designed to be administered by a trained clinician, and it contains 17 items rated on either a 3- or 5-point scale, with the sum of all items making up the total score. The HRSD may be a useful scale for cognitively impaired patients who have difficultly with self-report instruments.

5.2.2. Geriatric Depression Rating Scale

The 35-item Geriatric Depression Rating Scale (GDRS) combines the severity rating format of the HRSD with the content of the GDS. The GDRS probably requires less experience and training to administer than the HRSD and has good psychometric properties with hospitalized, outpatient, and community-dwelling older adults.

5.3. Instruments for Assessing Depression with Coexisting Dementia

Individuals with dementia pose a particular challenge when assessing for depression due in large part to the overlap of symptoms of depression and dementia as well as the effects of cognitive impairments and memory deficits. Several instruments have been developed to address the complicated nature of assessing depression with coexisting dementia, including the Cornell Scale for Depression in Dementia (CS) and the Dementia Mood Assessment Scale (DMAS).

5.3.1. Cornell Scale for Depression in Dementia

The CS contains 19 items that assess for the severity of depression based on a 3-point scale with content similar to the that of the HRSD. Psychometric properties of the CS are adequate. The CS appears to be a useful measure for assessing severity of depressive symptoms in older adults with dementia or cognitive impairments.

5.3.2. Dementia Mood Assessment Scale

The 17-item DMAS is a brief measure of mood in patients with dementia that includes a semistructured interview and direct observation of the individual over time. The psychometric properties of the DMAS are not strong. Its advantage lies in the combination of interview and direction observation.

In summary, there are a wide range of assessment instruments available for the assessment of depression among older adults, with each having strengths and weaknesses. Although many of the instruments were developed for use with younger adults, they have proven reliable and valid with older adults with a wide range of clinical characteristics.

6. TREATMENT OF LATE-LIFE DEPRESSION

Despite the availability of treatments for depressed older adults, only a few older adults utilize such treatments. Depressed older adults are most frequently first seen by a primary care physician, usually for somatic symptoms and complaints of low energy. Unfortunately, many older adults and their physicians attribute these symptoms to "normal aging" or a physical ailment and fail to make a mental health referral.

Although presentation and diagnostic considerations for depression appear to differ in older adults when compared to their younger counterparts, Karel and Hinrichsen's review of the literature suggests that treatment interventions effective for younger adults are also efficacious for older adults. This section focuses on psychosocial interventions, with an emphasis on evidence-based treatments for older adult depression.

6.1. Evidence-Based Psychosocial Treatments

A number of psychosocial interventions have proven effective for treatment of older adult depression. These treatments include psychodynamic, cognitive, behavioral, cognitive–behavioral, and interpersonal therapies. Although the rationale for each form of therapy is derived from varying conceptualizations, a number of studies have demonstrated the effectiveness of each of these therapies for treatment of depressed older adults.

6.1.1. Psychodynamic Therapies

Psychodynamic therapies focus on intrapsychic conflicts that may be affecting the individual's coping and adjustment. Emotional insight, gained through the therapeutic

process, is the goal of treatment for resolution of these conflicts and establishment of more effective coping styles. In addition, this form of therapy focuses on resolution and acceptance of social and physical losses that are characteristic of old age and addresses unresolved issues from various stages of development that may be contributing to the person's distress. Psychodynamic therapies also frequently incorporate a focus on the "self" and maintaining self-esteem through the trials and tribulations of the aging process.

6.1.2. Behavior Therapy

Behavior therapy relies on the premise that overt activity and mood are closely linked. In 1974, Lewinsohn, who offered one of the principal behavioral theories of depression, demonstrated that the lack of sufficient positive reinforcement from one's environment can negatively affect one's mood state. Zeiss and Lewinsohn adapted Lewinsohn's behavioral approach for older adults. This form of therapy typically involves behavioral activation, scheduling of pleasant events into the person's life, decreasing one's participation in aversive activities, and problem solving and social skills training. In addition, "homework" assignments are frequently given to facilitate the achievement of clearly developed goals and skills and activate the person for change.

6.1.3. Cognitive Therapy

Cognitive therapy, as developed by Beck *et al.* in 1979, assumes that maladaptive and irrational cognitions are contributing to, if not causing, the depressed mood. This therapy focuses on identifying distorted or irrational thoughts and beliefs that may be contributing to the mood state. During therapy, the person is encouraged to monitor, challenge, and eventually replace negativistic and irrational thoughts with realistic, more positive ones. Practice through completion of homework assignments is integral to this form of therapy, and clients are encouraged to monitor and challenge irrational or negativistic cognitions in their daily lives once they have practiced in a therapeutic setting.

6.1.4. Cognitive–Behavior Therapy

Gallagher-Thompson and Thompson combined behavioral and cognitive therapies, which they termed cognitive–behavior therapy, for the treatment of depression in older adults. This form of therapy is based on the rationale that irrational cognitions and/or maladaptive behaviors contribute to and maintain a depressed mood state. It is assumed that negative cognitions or thoughts, especially about one's self, future, and experiences, create a negative lens through which one sees the world. In addition, the depressed older adult is not engaging in enough positive mood-enhancing activities, which tend to exacerbate the negative cognitions. This form of therapy is generally psychoeducational, goal-directed, structured, and brief, designed to give the depressed older adult the cognitive and behavioral skills (e.g., relaxation techniques, assertiveness training, and problem-solving skills) he or she is lacking.

6.1.5. Interpersonal Psychotherapy

Finally, interpersonal psychotherapy (IPT), developed by Klerman *et al.* in 1984, is a time-limited form of therapy that emphasizes interpersonally relevant issues that may have preceded or resulted from depression. Specifically, this therapy stresses the importance of family relations in older adults' lives and, as such, emphasizes the nature of important relationships and focuses on one of four potential interpersonal problems: grief, interpersonal dispute, role transition, and interpersonal deficits. Treatment strategies are then implemented to address each of these problems and reach identified treatment goals. In a review of the use of IPT with older adults, Hinrichsen found several studies that demonstrated the effectiveness of this therapy for older depressed adults.

6.1.6. Other Psychosocial Treatments

Although the previously described therapies have been well established as effective for use with older adults, a number of other therapeutic approaches are used, although little empirical evidence is available for their efficacy in this population. It should be noted, however, that simply because an intervention has not been empirically studied does not indicate that it is not useful or effective in the treatment of older adult depression. Several of these forms of treatment have been examined through preliminary studies that have rendered promising results.

6.1.6.1. Life Review Therapy   Life review therapy, a form of reminiscence therapy, encourages the older adult to reflect on his or her life and is thought to facilitate his or her progress through Erikson's final developmental task and emerge with a sense of integrity

or despair. Although it is used for the treatment of depression in older adults, some caution its use with this population because it could prove more harmful than helpful, especially for older adults for whom a life review would not be a positive experience. Researchers continue to debate the efficacy of this form of therapy.

6.1.6.2. Group Therapy   Group therapy, especially psychoeducational group therapy, is another form of treatment for older adult depression. Group therapy captures the interpersonal or social component not seen in other forms of therapy and thus allows for older adults to share with and learn from each other while acquiring behavioral and cognitive coping skills. This form of therapy can be quite effective with depressed older adult family caregivers.

6.1.6.3. Less Common Forms of Treatment   Less common forms of treatment for older adult depression include bibliotherapy and self-help therapy, both of which emphasize the individual's role in treatment and overcoming depression without a great deal of intervention from a therapist. Typically, in both of these forms of therapy, the older adult is educating himself or herself through readings. In 1989, Scogin *et al.* conducted one of the few studies on the efficacy of bibliotherapy and found promising and generally positive results.

A number of treatments or therapies for older adult depression have been described briefly in this section. Although there are several well-established, effective treatments for depression in this population, there is a need for additional research to empirically examine some of the other less common but potentially advantageous forms of therapy. There is also a need to make treatments more available and acceptable to older adults so that when referred by physicians, depressed older adults will seek treatment.

See Also the Following Articles

Aging and Competency ■ Elder Caregiving ■ Psychotherapy in Older Adults

Further Reading

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Blazer, D. (2003). Depression in late life: Review and commentary. *Journal of Gerontology: Medical Sciences*, *58A*, 249–265.

Blazer, D. G. (2002). *Depression in late life*. New York: Springer.

Coon, D. W., Rider, K., Gallagher-Thompson, D., & Thompson, L. (1999). Cognitive–behavioral therapy for the treatment of late-life distress. In M. Duffy (Ed.), *Handbook of counseling and psychotherapy with older adults* (pp. 487–510). New York: Basic Books.

Edelstein, B. A., Kalish, K. D., Drozdick, L. W., & McKee, D. R. (1999). Assessment of depression and bereavement in older adults. In P. Lichtenberg (Ed.), *Handbook of assessment in clinical gerontology* (pp. 11–58). New York: Wiley.

Fiske, A., Kasl-Godley, J. E., & Gatz, M. (1998). Mood disorders in late life. In B. Edelstein (Ed.), *Clinical geropsychology* (Vol. 7, pp. 193–229). New York: Elsevier.

Frank, E., Frank, N., Cornes, C., Imber, S. D., Miller, M. D., Morris, S. M., & Reynolds, C. F. (1993). Interpersonal psychotherapy in the treatment of late-life depression. In G. L. Klerman, & M. M. Weismman (Eds.), *New applications of interpersonal psychotherapy* (pp. 167–198). Washington, DC: American Psychiatric Press.

Gallagher-Thompson, D., & Thompson, L. W. (1996). Applying cognitive–behavioral therapy to the psychological problems of later life. In S. H. Zarit, & B. G. Knight (Eds.), *A guide to psychotherapy and aging: Effective clinical interventions in a life-stage context* (pp. 61–82). Washington, DC: American Psychological Association.

Gatz, M., & Fiske, A. (2003). Aging women and depression. *Professional Psychology: Research and Practice*, *34*, 3–9.

Gatz, M., Fiske, A., Fox, L., Kaskie, B., Kasl-Godley, J., McCallum, T., & Wetherell, J. (1998). Empirically-validated psychological treatments for older adults. *Journal of Mental Health and Aging*, *41*, 9–46.

Gatz, M., & Hurwicz, M. (1990). Are old people more depressed? Cross-sectional data on Center for Epidemiological Studies Depression Scale factors. *Psychology and Aging*, *5*, 284–290.

Hinrichsen, G. (1999). Interpersonal psychotherapy for late-life depression. In M. Duffy (Ed.), *Handbook of counseling and psychotherapy with older adults* (pp. 470–486). New York: Wiley.

Jamison, C., & Scogin, F. R. (1992). Development of an interview-based geriatric depression rating scale. *International Journal of Aging and Human Development*, *35*, 193–204.

Karel, M., & Hinrichsen, G. (2000). Treatment of depression in late life: Psychotherapeutic interventions. *Clinical Psychology Review*, *20*, 707–729.

Karel, M. J., Ogland-Hand, S., & Gatz, M. (2002). *Assessing and treating late-life depression: A casebook and resource guide*. New York: Basic Books.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.

Knight, B. G., Santos, J., Teri, L., & Lawton, M. P. (1995). The development of training in clinical geropsychology. In B. G. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.),

Mental health services for older adults: Implications for training and practice in geropsychology. Washington, DC: American Psychological Association.

Mulsant, B. H., & Ganguli, M. (1999). Epidemiology and diagnosis of depression in late life. *Journal of Clinical Psychiatry*, *60*(Suppl. 20), 9–15.

Parmalee, P. A., Katz, I. R., & Lawton, M. Pl. (1993). Anxiety and its association with depression among institutionalized elderly. *American Journal of Geriatric Psychiatry*, *1*, 46–58.

Storandt, M., & VandenBos, G. R. (Eds.). (1994). *Neuropsychological assessment of dementia and depression in older adults: A clinician's guide*. Washington, DC: American Psychological Association.

Weissman, M. M., Bruce, M. L., Leaf, P. L., Florio, L. P., & Holzer, C., III. (1991). Affective disorders. In L. N. Robins, & D. A. Regier (Eds.), *Psychiatric disorders in America: The epidemiologic Catchment Area Study* (pp. 53–80). New York: Free Press.

Developmental Counseling

David A. Jepsen

University of Iowa, Iowa City, Iowa, USA

GLOSSARY

appraisal Collecting and synthesizing data to construct a description of a client's life history and current problems.

collaborative relationship A counselor and client working together jointly toward agreed-on goals.

developmental change Relatively enduring irreversible differences over time in a person's activity pattern judged as movement toward mastery of culturally valued functions and social competencies.

developmental problem A "gap" between present experiences and experiences desired by the developing person or by people who have a stake in the person's development.

interventions Counselor-initiated acts that literally "go between" events sequenced along a client's life span, between the client's problem-saturated activity and future growth/movement toward resolving problems and addressing new challenges.

interventions for prevention Interventions designed to prevent problems from either becoming more serious or expanding to other areas of functioning.

life cycle Differences in a person's activities across time.

life structure Differences in a person's activity pattern or configuration at a particular point in time.

Developmental counseling is a collaborative and catalytic relationship formed between a professional counselor, who is trained in both counseling and human development, and a client, who is currently experiencing or is expected to experience problems that impede progress toward mastering personally and/or culturally valued functions for the purpose of co-constructing new meanings of current actions and/or new possibilities for future actions.

1. INTRODUCTION

Most people experience problems while growing up, that is, problems that impede their progress toward mastering personally and/or culturally valued functions. Most contemporary cultures value successful functioning as a friend, a learner, a producer/worker, a spouse or partner, and a citizen. Furthermore, the specific cultural group with which a person identifies—not necessarily the dominant culture—sets a "timetable" for growing up that specifies the period in life during which the person's attentions to these functions should begin, peak,

and diminish. Problems arise when people perceive a "gap" between their current experiences and the experiences desired by themselves, by important others in their lives, and/or by cultural standards. People may experience developmental problems such as making and keeping friends, finding and sustaining meaningful work, achieving intimacy, and defining and implementing a sense of unique personal identity.

Described in another way, a person's development is problematic when his or her activity is not organized effectively to perform major life functions that the person's culture considers to be fundamental to the human condition. Although most cultures seem to agree on some essential human functions, they differ on others. For example, achieving a sense of personal identity is more prominent in cultures that emphasize individualism and autonomy than in those that emphasize community and belonging. Cultures also differ in their prescribed cultural pathways to achieving competence. For example, some cultures promote pathways of personal independence, whereas others promote pathways of social interdependence.

Over the past 50 years or so, several writers have advocated an approach to counseling that addresses developmental problems. The aggregate scholarship on developmental counseling remains somewhat informal and eclectic. Perhaps the most systematic treatment has been developmental counseling and therapy (DCT) advanced by Ivey and colleagues. Given such wide-ranging ideas and practices, this article represents a composite definition amalgamated from many sources.

Developmental counseling represents a broad counseling approach designed to help a wide range of people who are experiencing problems of growing up across the life span. Because few counselors can master all of the skills and knowledge necessary to serve all people with developmental problems, developmental counselors usually specialize in working with clients from a particular age group (e.g., counselors for children), address a particular cultural function (e.g., career counselors), or identify with a particular subculture (e.g., counselors for women).

The delivery of developmental counseling can be distinguished from other counseling approaches. Most notably, developmental counselors work with people to activate their movement toward developmental change rather than to remediate pathology. Developmental counselor–client contacts usually are for a limited number of meetings and may be intermittent over longer periods of time, similar to the relationships that many adults maintain with their dentists or accountants.

Developmental counselors serve as catalysts, rather than as the sole providers or vehicles, for clients to change and mobilize every helping resource in their environment—sometimes called activating a network of resources—so as to stimulate and support developmental changes. Client growth is expected to occur both outside and within the counseling relationship.

Developmental maturation or growing up is given meaning both from the perspective of the person who is experiencing it and from the perspective of external observers. If the person who is experiencing developmental problems seeks the help of an expert "external observer" called a professional counselor, that collaboration is called developmental counseling. Sometimes, developmental counselors will work with groups of people who are about to experience common developmental problems in an effort to prevent more serious problems. Professional counselors provide, among other initiatives, a fresh perspective on developmental problems experienced by clients, a perspective that is informed by knowledge of human development and cultural standards. Although developmental counselors employ many of the same ideas and techniques as do counselors with other theoretical preferences, the former are most clearly set apart by subscribing to a developmental viewpoint.

2. THE DEVELOPMENTAL VIEWPOINT

Development generally refers to changes in patterns of overt and covert activity as people grow up or mature in their capacity to perform several culturally valued life functions. Individual development involves multiple changes as the person strives to master many life functions simultaneously. When such changes are judged as movement toward cultural ideals of maturity or healthiness, there is general agreement that the person made developmental changes. Developmental change is progressive and incremental, usually involving movement away from simple, dependent, self-centered activity patterns that are typical of younger members of a society and movement toward complex, independent, society-centered activity patterns that are typical of mature members of a society. The metaphor for developmental change is usually a spiral rather than an inclined plane, thereby representing back-and-forth movement that is predominantly forward. By contrast, development is not construed as change in traits (e.g., becoming less

introverted or more intelligent), as change in states (e.g., becoming less anxious or more expressive), or as change in specific isolated behaviors (e.g., eliminating socially disruptive behaviors, increasing frequency of approaching peers).

Developmental counselors conceptualize differences in individual clients' activity along two dimensions: the time across the life cycle and the activities within a contemporary life structure. The life cycle traces differences in a person's activities across time, and the life structure specifies differences in a person's activity pattern or configuration at a particular point in time. Development usually connotes quantitative changes (e.g., frequency, intensity, duration) in age-graded normative activity that is typical, rather than unusual, for the age group. For example, the duration of friendships tends to lengthen during late childhood and adolescence. Development also refers to qualitative changes in action patterns (e.g., complexity, coherence) that serve the same broad, culturally valued life functions. For example, patterns of values expressed as reasons for moral or career choices are generally more complex and coherent during adolescence than during childhood. Developmental change is also reflected in the transformation of the organization of activity—the life structure—as well as in the activity content.

Life span simply denotes the sequence of activities across the whole of life from birth to death. In contrast, the developmental life cycle denotes changes that meet three conditions:

- A sequence of changes over the life span that are undergone as people master progressively more complex life functions
- Changes over time through segments of the life span, called life stages, that are characterized by distinctly different activities and expectations
- Stages of change that replicate the progressive mastery experienced by preceding generations

An essential idea in development is the life stage, an age-graded segment of the life cycle that is demarcated by changes in the physical body, in social expectations, and (consequently) in the psychological activity responding to these changes. Each life stage includes a set of developmental tasks that must be mastered, or normative problems that must be solved, within an age range as expected by the culture and, therefore, must follow the culturally imposed timetable. These tasks can be carried out well or poorly. Mastery of tasks within an early stage is essential to mastering other tasks within that stage and as a precursor to mastering tasks within later stages.

Failure to resolve tasks from an earlier stage often is experienced as unfinished business and is likely to produce recurrent problems later in life, sometimes called delays in development. Developmental counselors often encounter young adult clients who are experiencing developmental problems because they did not master the tasks during childhood and adolescence.

In contrast to developmental tasks are developmental marker events that occur at various ages/stages and that require the person to make adaptations in activities and roles. Such events often result from circumstances beyond the person's control (e.g., illness, death of a loved one, accident, job loss, military conscription, natural disaster). Other marker events are likely to be anticipated (e.g., marriage, childbirth, job changes, retirement).

Developmental life structure denotes the contemporary pattern or configuration of activities organized to fulfill culturally valued functions and to achieve age-appropriate developmental tasks. Activity patterns are often distinguished as social roles (e.g., friend, student, jobholder, citizen), each of which represents aggregate responses to cultural expectations and developmental tasks. The multiple social roles that one person plays often introduce problems such as conflicting demands from two or more roles. For example, a worker may be expected to meet production goals that require time and energy usually devoted to family. As people age and developmental tasks change, the activity that is considered appropriate for the age and role also changes; therefore, life structures are reorganized into new activity patterns. For example, when high school or college graduates move to different communities, start new jobs, and start families, their activity patterns become reorganized. The transition from one life structure to another presents a person with an opportunity to address new developmental tasks and problems.

3. COLLABORATIVE RELATIONSHIPS

The developmental counselor and client develop a working relationship, sometimes called a working alliance, that represents a collaboration between two perspectives on the client's experiences, the client's internal subjective perspective, and the counselor's external "objective" perspective. The collaboration recognizes that the counselor is influenced by the client and also exercises influence on the client; the client tends to talk and act in the manner in which the counselor

listens and responds. Their shared goal is to accelerate movement toward achieving mastery of developmental changes (e.g., expanding action possibilities within the current structures, reorganizing structures). Co-constructing a developmental goal is often a first step in the developmental counseling process. Developmental counseling goals emphasize constructing and not repairing, working toward new client successes in mastering life functions and not recovering the functions—or, in short, promoting healthy development rather than treating psychopathology.

4. GUIDING PRINCIPLES FOR DEVELOPMENTAL COUNSELING PROCESSES

Three principles guide the counseling processes selected by developmental counselors. A paramount principle is that counselors adapt their processes to fit clients' current developmental stages and structures. For example, developmental counselors who work with children will adapt their manner of communication as well as the content of their language to fit the children's cognitive developmental levels. Because people under extreme stress often regress to forms of communication appropriate for earlier stages (e.g., they cry or scream like children), counselors must be careful to initiate responses that are appropriate matches regardless of clients' chronological ages or current social standing.

Such counselor adaptation is sometimes called the "problem of the match" in developmental counseling or instruction. The principle was probably derived from education (rather than from psychology or medicine), where the phrase "developmentally appropriate practice" is frequently invoked to describe instructor selection of learning climates and pedagogical methods that match the students' cognitive developmental level. Some counselors are very specific in matching cognitive developmental structures (e.g., assessing for client Piagetian stages), whereas others match on the broad dimension of conceptual complexity. Developmental counselors believe that their effectiveness is linked to their ability to adapt counseling processes to match their clients' developmental levels.

A second principle guiding developmental counseling processes is that counselors' stances toward their clients cycle between support and challenge. Counselors begin by supporting their clients' full expression of problem-filled stories. Then, counselors challenge—without confronting—their clients to move toward actions that promote developmental growth. Later, counselors support their clients' efforts to think or act in new ways that will contribute to growth. Some developmental counselors attempt to challenge their clients at "one-half step" beyond the clients' current positions on a theoretical developmental continuum. The rationale is, in part, to present the next steps in solving the developmental problem as a way in which to encourage client changes.

Third, most developmental counselors adhere to a principle captured by the following phrase: "the symbiotic nature of appraisal and intervention." The principle holds that both processes are ongoing within the counseling process and that each contributes to and profits from the other. For example, appraisal data are the material for counselor or client interpretive interventions, and client responses to interpretations become data integrated into client portraits constructed from appraisal data.

5. APPRAISAL IN DEVELOPMENTAL COUNSELING

Appraisal in developmental counseling is the conjoint collecting and synthesizing of data about the client's life structure and life cycle. The data are used to construct a descriptive portrait of the client's development and current problems. The assumption is that current developmental problems are best understood when the person's contemporary activity patterns and the sequence of past events are fully examined. Appraisal of a client's life structure involves identifying patterns or configurations of contemporary roles, relationships, and/or activities. For example, adult clients are usually asked to describe the work, family, and community roles that they play, whereas young clients are often asked to describe their family roles and family compositions, sometimes in the form of a family tree. Specifically, the client and counselor review the life structure for incidents of both the support and stressors.

Life cycle appraisal involves describing the sequence of events that the client deems to be critical in his or her development. In addition, the developmental counselor inquires about how the client has experienced developmental tasks within past and current life stages. Life cycle appraisal processes may involve, for example, a life history questionnaire or a life review interview. The client's history is assumed to affect current experiencing.

Therefore, developmental counseling appraisal techniques emphasize the person's assets and coping strategies as well his or her limitations and barriers.

When making an appraisal, the developmental counselor gives careful attention to the context or surroundings of the client's activity. This attention often yields important observations about person–environment interactions that are critical to understanding the client's development. Questions about social and physical surroundings are frequently included in early discussions with the client. Parents, teachers, and/or friends may also be consulted for independent observations of the client's surroundings. Attention is paid to the hierarchical organization of social systems, especially the differential exercise of authority and power, and to the horizontal structure for opportunities for peer relationships. Specifically, the contextual barriers to action and opportunities for actions are identified.

6. DEVELOPMENTAL COUNSELING INTERVENTIONS FOR RESOLVING CURRENT DEVELOPMENTAL PROBLEMS

Interventions are counselor-initiated events that literally "go between" events sequenced along the life span, that is, between the client's current problem-saturated activity and his or her future growth or movement toward resolving problems and addressing new developmental tasks. Developmental interventions usually focus on the client's cognitions or behaviors and use techniques similar to other counseling approaches such as interpretations, teaching, and role-playing.

The developmental counselor frequently constructs interpretations and invites the client to construct interpretations that stimulate new client insights. These interpretations emphasize explanations derived inductively from descriptions rather than derived deductively from classification systems. The developmental counselor may also teach the client a new action sequence that is derived from the counselor's knowledge about the action sequences, called developmental pathways, which are common to people who have mastered a developmental task. Role-playing is another example of how the counselor teaches the client a new pattern of activity. The counselor demonstrates the activity, and the client imitates it, practices it, rehearses it, and finally implements it under real-life conditions. The developmental counselor also encourages the client to engage in self-directed activities (e.g., self-help books, computer-assisted programs) or to take advantage of learning opportunities outside the counseling sessions.

7. DEVELOPMENTAL COUNSELING INTERVENTIONS FOR PREVENTING OR REDUCING DEVELOPMENTAL PROBLEMS

The developmental counselor often engages in interventions to prevent people's problems either from becoming acute or from expanding to interfere with more life functions. Primary prevention occurs when the counselor engages in educating groups of people who are addressing a common developmental task. For example, the counselor may instruct groups of people who are preparing for retirement or, in developmental terms, are preparing to cope with the task of assigning meaning to their lifetimes of productive activity. The instructional goal is to reduce acute problems associated with the loss of work. By learning coping strategies and constructing action plans, these people may avoid persistent regrets or sadness about retiring.

Secondary prevention occurs when the developmental counselor works with a small group of people selected because the members have a common developmental problem. The goal is to prevent the spread or further complication of problematic actions or situations. For example, the developmental counselor may work with young adults who have few friends and are considered to be "shy" in an effort to enrich their current friendships and possibly increase the number of friends or acquaintances, thereby forestalling more serious relationship problems.

See Also the Following Articles

Lifespan Development and Culture

Further Reading

Arbona, C. (2000). The development of academic achievement in school aged children. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed.). New York: John Wiley.

Blustein, D. L., Juntunen, C. L., & Worthington, R. L. (2000). The school-to-work transition: Adjustment challenges of the forgotten half. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed.). New York: John Wiley.

Crespi, T. D., & Generali, M. M. (1995). Constructivist developmental theory and therapy: Implications for counseling students. *Adolescence*, *30*, 735–744.

Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology*, *54*, 461–490.

Hesketh, B. (2000). Prevention and development in the workplace. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed.). New York: John Wiley.

Hill, R. D., Thorn, B. L., & Packard, T. (2000). Counseling older adults: Theoretical and empirical issues in prevention and intervention. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed.). New York: John Wiley.

Howard, G., Nance, D., & Myers, P. (1986). Adaptive counseling and psychotherapy: An integrative eclectic model. *The Counseling Psychologist*, *14*, 363–442.

Ivey, A. E. (1986). *Developmental therapy: Theory into practice*. San Francisco: Jossey–Bass.

Ivey, A. E., & Goncalves, O. F. (1988). Developmental therapy: Integrating developmental processes into the clinical practice. *Journal of Counseling and Development*, *66*, 406–413.

Juntunen, C. L. (2002). Development, developmental concerns, and counseling. In C. L. Juntunen, & D. R. Atkinson (Eds.). (2002). *Counseling across the lifespan: Prevention and treatment*. Thousand Oaks, CA: Sage.

Juntunen, C. L., & Atkinson, D. R. (Eds.). (2002). *Counseling across the lifespan: Prevention and treatment*. Thousand Oaks, CA: Sage.

Vera, E. M., & Reese, L. E. (2000). Preventive interventions with school-age youth. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed.). New York: John Wiley.

Diagnostic and Statistical Manual of Mental Disorders

Victoria del Barrio

Universidad Nacional de Educación a Distancia, Madrid, Spain

GLOSSARY

axis A reference line in a coordinate system.
category Each division in a system of classification.
classification Grouping of elements according to classes.
code Number for identification of disorders.
mental disorder A clinically behavioral or psychological syndrome associated with stress or disability or with increased risk of suffering death, pain, or an important loss of freedom.
syndrome A group of associated disease symptoms.

The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) classification system for mental disorders, developed by the American Psychiatric Association and now employed worldwide, has its historical roots in previous systems dating back over several centuries, from the Greek Hippocrates in the fourth-century BC to the nineteenth-century German psychiatrist Emil Kraepelin. This article traces the development of the *DSM*, describing successive changes and improvements over the course of its six editions, as well as its relationship with the other main system, the European-based *International Classification of Diseases*.

1. ANTECEDENTS

Classification is both the process and the result of arranging individuals into groups formed on the basis of common characteristics. It is important in any science, but it is usually a difficult task that raises a variety of problems for experts to deal with, including mixture (heterogeneity or homogeneity of a group), discrimination (between-group differences), and identification (assignment of an individual to a group).

Historically, the first attempt to put the variety of psychological disorders into some kind of order was made by the Greek physician Hippocrates (4th century BC); he coined the terms *phrenesis*, *mania*, and *melancholia*, which were maintained by Galen in Rome (1st century AD). Much later, during the Renaissance, Barrough (1583) divided mental disorders into mania, melancholia, and dementia. The work was continued in Europe by the German philosopher Immanuel Kant in his *Anthropologie* and by the Frenchman F. Pinel in his *Nosologie Philosophique* (1789), which divided mental

disorders into mania with and without delirium, melancholia, dementia, and idiocy. The Swede, K. Linnaeus, better known for his classifications of plants and animals, also extended his work to the field of the human mind, employing such categories as Ideales (delirium, amentia, mania, melancholia, and vesania), Imaginnarii (hypochondria, phobia, somnambulism, and vertigo), and Patheci (bulimia polydypsia, satyriasis, and erotomania).

At the end of the 19th century, the German, Emil Kraepelin—often considered the founder of modern scientific psychiatry—in his *Lehrbuch der Psychiatrie* [*Handbook of Psychiatry*] (1899) included a highly influential classification. He aimed to identify groups of patients through symptom clusters that would define different syndromes, and he attributed disorders to four organic roots: heredity, metabolic, endocrine, and brain disease. His categorization laid the bases for systems such as the first official American Psychiatric Association (APA) classification, which included mental disorders by brain traumatism, mental disorders by encephalopathy, mental disorders by drugs, mental disorders by infectious agents, syphilis, arteriosclerosis, epilepsy, schizophrenia, manic–depressive psychosis, psychopathy, psychic reaction, paranoia, oligophrenia, and other cases.

In the 1960s, neo-Kraepelinian psychiatrists called for improvements to the old classification, and men such as Spitzer, Endicott, and Feighner began to prepare the construction of a new American psychiatric classification.

Although the roots of modern American classification are in Europe, the first official classification was a census of mental disorders drawn up in the United States in 1840, with just two categories: idiocy and madness. Another important antecedent of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) was the work of the Committee on Statistics of the American Psychiatric Association (1917), led by Dr. T. Salmon, which recorded mental disease statistics from the psychiatric services of hospitals. Its result was the *Statistical Manual for the Use of Hospitals for Mental Disease*. A few years later, the New York Academy of Medicine organized the National Conference on Disease Nomenclature (1928) to unify the terminology and nosology used by practitioners; its *Standard Classified Nomenclature of Diseases* (1932; rev. ed., 1934) originally referred only to physical diseases, but a section on Diseases of the Psychobiological Unit was included later, and the 1951 edition was published for use in mental hospitals. All of these classifications owed a debt to Kraepelin's system.

World War II brought so many diagnostic problems to military psychiatrists that they urged the construction of a *Classification of Mental Problems*. The psychological nomenclature was introduced in 1944 by the Navy and a year later by the Army. However, the existing *Standard Nomenclature* (1932) proved inadequate for many situations, and the Armed Forces asked for a revision of the three systems in use. The result was the creation of the *DSM*s.

In Europe, a parallel movement, that of mental hygiene, held its first meeting in Paris in 1932. A European classification would unify psychiatric nomenclature, with the following categories: congenital mental disorders, endogenous psychosis, reaction psychosis, symptomatic psychosis, encephalopathies, epilepsy, toxicomany, and non-aliened mental disorders. However, this proposal was not well accepted by the mental health community.

Today, the two most well-known categorical classifications of mental disorders are the *International Classification of Diseases* (*ICD*) and the *DSM*. The *ICD*, created in Europe, has been promoted by the World Health Organization (WHO) since 1900, but its original version included only physical diseases; the *DSM*, published by the APA since 1952, has always focused on mental disorders. The two classifications attempt to deal with the same behaviors and coincide to a large extent, but there are some important differences.

The first *ICD* versions included only physical disorders (Table I), but the *ICD-6* (1948) saw the incorporation of mental diseases, bringing it closer to the *DSM-I*. The *ICD-8* (1967) and *DSM-II* (1968) also coincided in many aspects, but the APA's decision to develop the *DSM-III* generated new tensions. Much of this divergence was resolved through better communication, and the *DSM-IV* and *ICD-10* reestablished the close relationship between the WHO and the APA.

2. DSM-I (1952)

The *DSM* was first published in 1952 (*DSM-I*). Influenced by its European antecedents, and by the views of A. Meyer (a biologically oriented psychiatrist) and K. Menninger (who was psychoanalytically oriented), it took into account not only biological but also social and psychological elements, and it offered a multidimensional consideration of disorders. The process of developing the manual involved the collaboration of the National Institute of Mental Health and different psychiatrists and military professionals.

TABLE I
Relationship over Time between the *DSM* and the *ICD*

Year	DSM (APA), United States	ICD (WHO), Europe
1900		ICD-1[a]
1910		ICD-2[a]
1920		ICD-3[a]
1929		ICD-4[a]
1938		ICD-5[a]
1948		ICD-6
1952	DSM-I	
1955		ICD-7
1966		
1967		ICD-8
1968	DSM-II	
1975		ICD-9
1978		ICD-9-CM
1980	DSM-III	
1987	DSM-III-R	
1992		ICD-10
1994	DSM-IV	
1996		ICD-10 Children
1998	DSM-IV-R	
2000	DSM-IV-TR	

[a]ICD with only physical diseases.

Its impact was quite limited. Smith and Fonda, in a study using the *DSM-I* criteria, found interrater agreement to be high for organic psychosis, melancholia, and schizophrenia but quite low in many other categories.

Other studies were also critical. For example, the omission of criteria was criticized by H. J. Eysenck and R. B. Cattell; in their view, common nomenclature in itself would not improve clinical diagnosis, as clinicians would understand different things from the same labels. Furthermore, P. E. Meehl questioned the reliability of clinical judgment, whereas L. Cronbach and Meehl drew attention to the urgent need for accuracy in diagnostic identification.

With the publication of the *ICD-7*, some of the differences between the WHO and APA classifications were highlighted. The *DSM-I* categories were also adopted for a computer program designed for assessment tasks (DIAGNO), but this presented some problems of application. The need to improve the system was evident, and it was Endicott, Guze, Klein, Robins, and Winokur who assumed the task of reviewing and improving the manual.

3. DSM-II (1968)

This was published as the *American National Glossary* to the *ICD-8*. An APA committee, with the help of some experts, carried out an in-depth revision of its first version. Meyer's previous influence was drastically reduced, whereas more room was given to psychoanalytical and Kraepelinian ideas in order to improve its acceptance by clinicians. It was based on the medical illness model of different syndromes made up of clinical symptom clusters.

Many unsuccessful attempts were made to coordinate the *DSM* and *ICD* systems, but the differences eventually increased, both in diagnoses and in terminology. For instance, the *DSM-II* had 39 more diagnoses than the *ICD-8 10/3*, and regarding terminology, terms associated with certain theoretical frameworks (such as "schizophrenic reaction" for schizophrenia) were avoided but others, such as "neuroses" or "psychophysiological disorders," remained.

The *DSM-II* has frequently been referred to as "old wine in a new bottle" since the influences of Meyer and Kraepelin were still present but many of the new research perspectives were absent. One year after its publication, Jackson described it as "a hotchpotch of different bases of classification." However, studies on schizophrenia diagnosis in various countries greatly stimulated the task of revision, and efforts were redoubled to clarify basic diagnostic criteria, paving the way for the construction of the *DSM-III*.

4. DSM-III (1980)

Criticism of the previous editions had noted the lack of an organizing criterion and the overlapping of categories. Efforts to improve the instrument culminated in the publication of the *DSM-III*.

A task force chaired by R. Spitzer, with international cooperation and the participation of N. Andreasen, J. Endicott, D. F. Klein, M. Kramer, Th. Millon, H. Pinsker, G. Saslow, and R. Woodruff, prepared the work. New members were added subsequently, and a draft was prepared at the important St. Louis conference in 1976. Collaborators were recruited from many institutions, including the Academy of Psychiatry and the Law, the American Academy of Child Psychiatry, the American Academy of Psychoanalysis, the American Association of Chairmen of Departments of Psychiatry, the American College Health Association, the American

Orthopsychiatry Association, and the American Psychological Association. Misunderstandings arose between the American Psychological Association and the American Psychoanalytical Association, but all were eventually solved.

The goals of these teams were (i) to expand the classification, maximizing its utility for the outpatient population; (ii) to differentiate levels of severity and cause within syndromes; (iii) to maintain compatibility with the *ICD-9*; (iv) to establish diagnostic criteria on empirical bases; and (v) to evaluate concerns and criticisms submitted by professional and patient representatives. Subsequently (1977–1979), drafts were sent to experts and practitioners for study and suggestions before the definitive modifications were made. The major characteristics of the *DSM-III* are (i) operational or explicit criteria for assigning patients to diagnostic categories and (ii) the implementation of a multiaxial framework.

4.1. Explicit Diagnostic Operational Criteria

Diagnostic criteria make explicit the signs and symptoms required for a diagnosis on the basis of empirically validated rules, and they are not only descriptive but also phenomenological and behavioral. They would permit greater effectiveness and reliability in diagnoses. Also included for each disorder was its description, the usual age at which it begins, mean duration, prognosis, rates by sex, and risk factors. Criteria for Axes I and II are basically the same as the Feighner Criteria, based only on disorder definitions, as in the *DSM-I* and *DSM-II*.

Such criteria were used when creating the *Research Diagnostic Criteria*, which included 23 disorders (schizophrenia, schizoaffective disorder–manic, schizoaffective disorder–depressed, depressive syndrome superimposed on residual schizophrenia, manic disorder, hypomanic disorder, bipolar with mania, bipolar with hypomania, major depressive disorder, minor depressive disorder, intermittent depressive disorder, panic disorder, generalized anxiety disorder, cyclothymic personality, labile personality, Briquet's disorder, antisocial personality, alcoholism, drug-use disorder, obsessive–compulsive disorder, phobic disorder, unspecified functional psychosis, and other psychiatric disorders).

An operative description of clinical disorders was included, employed in Axes I and II of the *DSM-III* to differentiate between syndromes. Unlike those of the biological classifications, *DSM-III* criteria are polythetic (patients do not need to have all the characteristics in order to be included in a category) and intentional (patient's characteristics are listed).

4.2. Inclusion of Five Diagnostic Axes

A new feature in the *DSM-III* was the inclusion of five diagnostic axes in order to obtain greater sensitivity and accuracy in diagnoses and treatments and permit the consideration of psychosocial aspects. The idea of such axes was introduced by Essen, Möller, and Wohlfahrt (in Sweden in 1947). They proposed separating the syndromes from all etiological conceptions, creating two principal axes ("phenomenological" and "etiological") plus three others ("temporality," "social functioning," and "others"). In 1969, Rutter *et al.* in the United Kingdom constructed a multiaxial classification system for children (with axes for clinical syndromes, delay of development, mental retardation, medical condition, and psychosocial situation), and other multiaxial systems were proposed in the United Kingdom by Wing in 1970, in Germany by Helmchen in 1975, and in the United States by Strauss. Apart from the *DSM-III*, the multiaxial formula was applied to the *ICD* in 1992.

Axes were assigned as follows: Axis I to clinical syndromes; Axis II (with two sections) to child maturational problems and adult personality disorders; Axis III to physically rooted problems; Axis IV to the intensity and severity of psychological stressors; and Axis V to the patient's level of adaptive functioning in the past year, objectively evaluated. Axes II, IV, and V link the syndromes with environmental determinants and reflect a shift toward more psychological conceptions.

DSM-III constitutes an authentic work of description on psychopathology (including differential diagnosis, etiology, treatment, prognosis, and management). The *DSM-III*, although atheoretical in principle, maintained a biological interpretation of the field. Although praised for its achievements, it was also criticized for including inessential characteristics of disorders and for employing too many diagnostic categories (265); psychoanalysts also criticized the lack of an axis that took into account defense mechanisms and ego functions, and they complained that the replacement of the term neuroses by anxiety disorders situated them at a greater distance.

In sum, the *DSM-III* represents a key stage in the evolution of the *DSM* classification. This change was both quantitative (from 104 syndromes to 395, and

from 132 pages to 500) and qualitative. Since 1980, with the adoption of the axes system, the corpus of categories has remained practically constant. Moreover, *DSM-III* made important progress in relation to aspects such as reliability, intracategory coherence and differentiation, validity, and behavioral data. It also favored the construction of new assessment instruments based on its criteria (self-rating questionnaires, semistructured interviews, and so on). Despite all these innovations and improvements, further revision was necessary, and this arrived in the form of the *DSM-III-R*.

5. DSM-III-R (1987)

This appeared in 1987 as the final result of revisions carried out by a work group headed by Spitzer. Innovations included the reorganization of categories; some improvements in Axes IV (on psychosocial stress) and V [inclusion of the Global Assessment of Relational Functioning (GARF) scale]; and the incorporation of aspects related to drug abuse, homosexuality (now included in unspecified sexual disorders), and hyperactivity. Classification of affective disorders was reorganized. This version had a dramatic impact worldwide and became more commonly used than the *ICD*, even in Europe.

6. DSM-IV (1994)

A task force led by Allen Frances working with experts and scientific groups from all over the world prepared this version. The goal was to improve its cultural sensitivity and to improve compatibility with the *ICD*.

Features of the new instrument included (i) brevity of criteria set, (ii) clarity of language, (iii) explicit statements of its constructs, (iv) it was based on up-to-date empirical data, and (v) better coordination with the *ICD-10*. The *DSM-IV* also presented a series of changes with respect to its predecessor. Categories such as organic mental disorders disappeared, whereas others such as eating disorders, delirium, dementia, and amnesic and other cognitive disorders, and some severe developmental disorders (Rett's syndrome and Asperger's syndrome), were incorporated; the categories of child and sexual disorders were restructured.

In the multiaxial system, the changes occurred in Axis IV, which includes more stress-generating events, and Axis V, in which other scales were added for determining level of adaptation—the Social and

Occupational Functioning Assessment Scale and the GARF scale.

Criticism of the *DSM-IV* has not been lacking, especially with regard to its focus on consensus more than on data. However (and despite claims that there would never be a *DSM-V*), remodeling of the *DSM-IV* is currently in progress.

7. CATEGORICAL DEVELOPMENT OF THE DSM

7.1. Important Changes

Over time, the *DSM* categories have undergone significant changes. Whereas in 1952 there were two main categories, in 2000 there were 17. This increasing complexity is even more evident in the case of sub-categories (Table II).

Each diagnostic category is identified by a numerical code. The current *DSM* codes the different disorders with its own code and also with that of the *ICD*, making it possible to compare and contrast diagnoses made with these two different categorical systems, thus facilitating communication between them.

7.2. Critical Positions

The *DSM* classification system has attracted criticism from various theoretical points of view. Those from the antipsychiatry current have always been against the use of classifications in psychiatry since they consider labeling a dangerous procedure. Another notable critic was Eysenck, who spoke of the fundamental weakness of any scheme "based on democratic voting procedures rather than on scientific evidence."

Further criticisms of the *DSM* refer to (i) the cultural biases of all classification systems; (ii) its extreme individualism—only individual diagnosis is taken into account; (iii) the influence on it of old-fashioned medical classifications, despite modern developments in psychology and psychiatry; and (iv) the "softness" of the categories, which are based more on description than explanation, despite the crucial importance of the latter.

However, despite such criticisms, categorical classifications have made possible comparisons and inferences that have helped to advance clinical knowledge. They may have their imperfections, but they have made an enormous contribution to diagnostic reliability and to understanding among mental health professionals. They have also helped to promote the creation of more

TABLE II
Evolution of *DSM* Categories

DSM-I (1952)	DSM-II (1968)	DSM-III (1980)	DSM-III-R (1987)	DSM-IV (1994)	DSM-IV-TR (2000)
B-Mental deficiency (6 diagnoses)	I-Mental retardation (6 diagnoses)				
Transient and situational personality disorders (8 diagnoses)	IX-Behavioral disorders of childhood and adolescence (7 diagnoses)	Disorders usually first evident in infancy and adolescence (45 diagnoses)	Disorders usually first diagnosed in infancy and adolescence (46 diagnoses)	Disorders usually first diagnosed in infancy and adolescence (46 diagnoses)	Disorders usually first diagnosed in infancy and adolescence (48 diagnoses)
			Delirium, dementia (1)	Delirium, dementia, and amnesic and other cognitive disorders (33)	Delirium, dementia, and amnesic and other cognitive disorders (29)
A-Organogenic disorders 1-Acute brain disorders (13 diagnoses)	II A-Psychoses associated with organic brain syndrome (40 diagnoses)	Organic mental disorders (58)	Mental disorders due to a general medical condition not elsewhere classified (61)	Mental disorders due to a general medical condition not elsewhere classified (3)	Mental disorders due to a general medical condition not elsewhere classified (3)
2-Chronic brain disorders (25)	II B-Non psychotic organic brain syndrome				
		Substance use disorders (119)	Substance-related disorders (22)	Substance-related disorders (133)	Substance-related disorders (140)
C 1-Psychotic disorders (17)	III-Psychoses not attributed to physical conditions listed previously (24)	Schizophrenic disorders (5)	Schizophrenia disorders (5)	Schizophrenia and other psychotic disorders (16)	Schizophrenia and other psychotic disorders (20)
		Paranoid disorders and other psychotic disorders (8)	Other psychotic disorders (5)		
		Affective disorders (9)	Mood disorders (9)	Mood disorders (16)	Mood disorders (16)
C3-Psychoneurotic disorders	IV-Neuroses (11 diagnoses)	Anxiety disorders (10)	Anxiety disorders (9)	Anxiety disorders (12)	Anxiety disorders (12)

C 2-Psychophysiologic autonomic and visceral disorders (10)	VI-Psychophysiological disorders (10)	Somatoform disorders (5)	Somatoform disorders (7)	Somatoform disorders (10)	Somatoform disorders (10)
		Factitious disorders (3)	Factitious disorders (3)	Factitious disorders (2)	Factitious disorders (2)
		Dissociative disorders (5)	Dissociative disorders (5)	Dissociative disorders (5)	Dissociative disorders (5)
		Psychosexual disorders (22)	Sexual disorders (20)	Sexual and gender identity disorders (31)	Sexual and gender identity disorders (31)
	VII-Special symptoms (10)			Eating disorders (3)	Eating disorders (3)
			Sleep disorders (12)	Sleep disorders (17)	Sleep disorders (17)
		Disorders of impulse control not elsewhere classified (6)	Impulse-control disorders (6)	Impulse-control disorders not elsewhere classified (6)	Impulse-control disorders (6)
		Psychological factors affecting physical condition (1)			
	VIII-Transient situational disturbances (5)	Adjustment disorders (8)	Adjustment disorders (9)	Adjustment disorders (6)	Adjustment disorders (6)
C 4-Personality disorders (18)	V-Personality disorders (33)	Personality disorders (12)	Personality disorders (12)	Personality disorders (11)	Personality disorders (11)
	X-Condition without manifest psychiatric disorders and nonspecific conditions	V-Codes for conditions not attributable to mental disorders that are a focus of attention or treatment	Other conditions that may be focus in clinical attention (9)	Other conditions that may be focus in clinical attention (32)	Other conditions that may be focus in clinical attention (36)
	XI-Non diagnostic term for administrative use		Other psychological factors affecting physical condition (1)		
		Additional codes	Additional codes	Additional codes	Additional codes
Total no. of syndromes 97	146	316	242	382	390

homogeneous and accurate assessment instruments that will constitute the source of future progress in the field.

See Also the Following Articles

Depression in Late Life ■ Forensic Mental Health Assessment ■ Mental Measurement and Culture

Further Reading

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 529, 281–203.

Eron, L. D. (1966). *Classification of behavior disorders.* Chicago: Aldine–Atherton.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Muñoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26, 57–63.

Haynes, S. N., & O'Brien, W. H. (1988). The Gordian knot of *DSM-III* use: Interacting principles of behavioral classification and complex causal models. *Behavioral Assessment*, 10, 95–105.

Millon, Th. (1996). The *DSM-III*: Some historical and substantive reflections. In Th. Millon (Ed.), *Personality and psychopathology*. New York: Wiley.

Schmith, H. O., & Fonda, C. (1952). The reliability of psychiatry diagnosis. *Journal of Abnormal and Social Psychology*, 52, 262–267.

Spitzer, R. L., Endicott, J., & Robins, E. (1975). *Research diagnostic criteria (RDC)*. New York: New York State Psychiatric Institute.

Spitzer, R. L., Gibbson, M., Skoldol, A. E., Williams, J., & First, M. B. (1994). *DSM-IV Casebook*: A learning comparison to the *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: American Psychiatric Press.

Spitzer, R. L., & Willson, P. T. (1975). Nosology and the official psychiatry nomenclature. In A. M. Freeman, H. I. Kaplan, & B. J. Sadock (Eds.), *Compressive textbook of psychiatry* (Vol. 2). Baltimore: Williams & Wilkins.

Szasz, T. (1966). The psychiatry classification of behavior: A strategy of personal constraint. In L. D. Eron (Ed.), *The classification of behavior disorders* (pp. 123–170). Chicago: Aldine.

Widiger, T. A., Frances, A. J., Pincus, H. A., First, M. B., Ross, R., & Davis, W. (1994). *DSM-IV sourcebook.* Washington, DC: American Psychiatric Association.

Diverse Cultures, Dealing with Children and Families from

Antoinette Halsell Miranda

Ohio State University, Columbus, Ohio, USA

race A social construct that attempts to divide human groups based on physical characteristics.

racism A belief that people can be validly grouped on the basis of biological traits and that some groups have superior traits, whereas others have inferior ones.

structural assimilation A process that occurs when one group can access and experience full participation within majority group organizations such as country clubs and private clubs.

GLOSSARY

cultural assimilation A process that occurs when one group assumes the values, behaviors, and beliefs of another group, often the majority culture.

culture The beliefs, values, traditions, behavior patterns, and social and political relationships that are shared by a group of people who are bound together by things such as common history, language, social class, and religion.

discrimination Often based on prejudice, a process consisting of differential treatment directed toward certain groups.

ethnicity An individual's sense of identification that provides him or her with a sense of belonging to a group.

power An important concept when entering into a discussion about diversity given that the struggle for power, and the dominant group having power, has played an integral part in the continued oppression of marginalized groups.

prejudice Attitudes and beliefs toward a particular group that are often negative and formed in disregard of the facts.

Working with culturally different children and families requires psychologists to have an understanding of cultural differences and, in the best of circumstances, to become cross-culturally competent.

1. INTRODUCTION

It is fairly well known that there has been a dramatic demographic shift in the United States population during the past decade or so. Based on the 2000 census, people of color make up approximately 29% of the U.S. population. By the year 2050, they are expected to account for 47% of the U.S. population. The demographic changes have major implications for many institutions, including schools and the workforce. The change in the ethnic composition of the population is viewed as both an opportunity and a challenge. There is the opportunity

to enrich society with multiple and diverse cultural elements. The challenge is to find effective ways in which to incorporate and respect the multiple diversities that exist while at the same time recognizing that there is an American culture. Historically, White ethnic groups have been "forced" to abandon their cultures and blend into the American culture. Ethnic minorities, on the other hand, have not been so successful. The reason why has not been for lack of trying; rather, they have not fared as well due to the racism and prejudice that exists in the United States. Ethnic minorities are often received and treated as marginalized groups. As marginalized groups, they have not always had full access to the "American dream." Although the civil rights movement and subsequent civil rights laws in the United States improved opportunities in education and employment for many ethnic minorities and women, people continue to be marginalized and discriminated against.

While many ethnic minorities have been successful in culturally assimilating, they still struggle to experience full structural assimilation in the United States. There has been a growing acceptance and recognition that, as a society, there has not been an appreciation of the rich diversity that exists; instead, there has been an attitude that one must fit in or assimilate to be successful. Exploring how variables such as class, race, ethnicity, and gender interact and intersect to influence behavior has been critical to researchers understanding life in America for marginalized groups. The United States is a pluralistic nation made up of many different ethnic groups whose members desire to retain their cultural identities. Understanding these different cultural identities is beneficial in working effectively with diverse populations. Identity development research has contributed to this knowledge base.

2. RACIAL AND ETHNIC IDENTITY

Identity is complex and influenced by a number of factors. Erickson was instrumental in helping people to understand that individual identity is embedded in social, cultural, and historical contexts. Race and ethnic identity adds to this complexity. The development of racial and ethnic identity for people of color often begins to occur during late adolescence. It is generally this period of time when adolescents start to explore the questions of "Who am I?" and "Who will I become?" Searching for racial and ethnic identity can be a lifelong journey. A focus on certain aspects of people's identities may vary at different times in their lives, depending on

the circumstances. For example, high school is a time when most adolescents begin to start to develop their identities separate from their families as they self-reflect on who they are. Individuals are faced with attempting to integrate these different components of identity to form who they are. Many aspects of society influence how people view themselves and, consequently, how they develop their identities.

Cross and Helms have researched Black identity and White identity, respectively. They have demonstrated that racial and ethnic identity develops in stages. Most important, these stages are fluid, with people moving back and forth between stages. It should also be noted that not all people move through these stages. Black adolescents and other adolescents of color tend to begin to develop racial identities earlier than do White adolescents because their self-perceptions can be influenced by how society views them. The aspect of people's identity that others notice and make them most aware of is one that can potentially loom large in their lives. Thus, adolescents of color are subjected to questions and stereotypes about their race and/or ethnicity more often than are White adolescents. Even though some adolescents begin to explore their racial and ethnic identities, research has shown that the movement through the stages most often occurs during late adolescence and early adulthood. For Whites, identity development involves the abandonment of entitlement or White privilege. The key challenge for people of color is the development of a positive identity. Research continues on the development of racial and ethnic identity.

3. UNDERSTANDING CULTURE AS A CONTINUUM

When working with culturally diverse children and families, it is important to understand that culture occurs on a continuum and is influenced by the process of dual socialization, which is the process by which individuals learn how to function in two distinct sociocultural environments. Most ethnic minorities in the United States are able to function in their own cultures but also learn how to navigate the waters of the majority culture. Individuals who navigate these two different cultures successfully are said to be bicultural. This concept is important and is related to assimilation. In the United States, the dominant group controls most of the social, economic, and political institutions. Thus, the dominant group determines these institutions'

norms, values, and beliefs, which are often referred to as culture capital. Having knowledge of culture capital allows easier movement through mainstream culture. Members of ethnic minority groups often must acquire the necessary cultural traits of the majority group to move up the social and economic ladders. It is often necessary for working-class people and people of color to learn the cultural norms, or to be told explicitly the rules of the culture of the dominant group, for them to acquire power.

However, when working with culturally different people, it is important to remember that individuals will vary along the cultural continuum. Individuals may choose to remain isolated from the mainstream, on one end of the continuum, or choose to relate to both cultures, never quite integrating them into one. Many ethnic minorities find themselves straddling two cultures. The challenge for these individuals is to develop strong self-concepts and positive self-esteem without compromising who they are. Understanding issues of identity and cultural continuum provides a better understanding of culturally different people. Thus, with the acknowledgment that diversity exists, a body of knowledge that encourages people to learn to respect, value, and understand diverse groups has evolved. This has occurred primarily through educational channels such as ensuring that disciplines require a diversity course for graduation and the development of multicultural competencies.

4. CROSS-CULTURAL COMPETENCE

As with so many disciplines, the demographics of education and psychology do not match those of the U.S. population. In psychology and education literature, there is increasing emphasis placed on training individuals to become cross-culturally competent. Cross-cultural competence means that individuals possess attitudes, abilities, and understandings that allow them to function not only in American culture but also within and across various ethnic, cultural, and linguistic groups. Three common elements of cross-cultural competence are personal awareness, knowledge, and skills. Although different disciplines may advocate diverse ways of attaining cross-cultural competence, these three elements appear to be universal aspects of that process. In education and psychology, having cross-culturally competent individuals is seen as critical to being able to deliver services that are sensitive and effective in meeting the needs of culturally different individuals. Developing cross-cultural competence is a

process. It is generally recognized that becoming cross-culturally competent will not occur by simply studying groups of people. It is necessary to have personal experiences with culturally different groups as well.

Self-awareness consists of individuals engaging in self-exploration that requires them to examine their own culture and socialization process that influences how they view and interact with majority and minority people. Understanding their own culture and its influence on their socialization is the first step in awareness. To understand another culture, individuals must first explore their own culture and its impact on their development. This often requires them to become aware of their biases and prejudices. Knowledge is the second component of cross-cultural competence. This is usually accomplished by studying the concept of diversity as well as racial and ethnic groups. This can be accomplished in a number of different ways, including reading books, talking with individuals from the cultural groups, and interacting with culturally different people. Skills are the final component. Skills involve the ability to work effectively with culturally diverse populations, understanding the influence that issues of diversity may have on people as well as one's ability to interact effectively with them.

Cross-cultural competence also means that individuals are aware of their own cultural limitations. Learning about other cultures requires them to take risks and have a willingness to be open, appreciative, and respectful of cultural differences. Intercultural interactions should be viewed as learning opportunities, recognizing that there is integrity and value in every culture one encounters.

5. CROSS-CULTURAL COMMUNICATION

One of the most difficult tasks that people face is communicating across differences. There is always the fear that what people intend to say will not be interpreted the way in which they meant it to be interpreted. It is generally recognized that there are high-context cultures and low-context cultures, and there are differences in how individuals from these cultures engage in communication. High-context cultures tend to communicate with fewer words and less emphasis placed on verbal interactions. As a result, individuals from these cultures are better at reading nonverbal cues. Situational and nonverbal cues convey primary meaning. For example, it is not unusual to hear an African American say, "All

my mother needed to give us was that "look." She didn't have to say anything. We knew we were in trouble." Individuals from high-context cultures also place a great deal of emphasis on establishing trust first. Thus, before revealing personal information, individuals from these cultures will want to first establish a relationship. Examples of high-context cultures include Asians, Native Americans, Arabs, Latinos, and African Americans. Low-context cultures, in contrast, tend to exhibit precise and direct verbal communication. They often do not notice, or are not attuned to, nonverbal communication. Meaning in low-context cultures is conveyed primarily through written and spoken words. Communication is often linear. Individuals from these cultures become particularly frustrated with high-context cultures that communicate in a circular fashion. Content of communication in low-context cultures can be best described as "cutting to the chase." Examples of low-context cultures include European Americans, Swiss, and Germans.

Cultures also differ in other aspects of communication, including proximity and touching. Groups differ in the amount of social distance they tolerate in social situations. For example, European Americans tend to keep a distance of 3 feet when talking to people with whom they are unfamiliar. Latinos and African Americans tend to feel comfortable with less social distance between themselves and others. Touching also varies greatly between groups. For example, handshaking takes on different meanings across cultures. Americans use handshaking as an introduction, whereas other cultures may be reluctant to engage in handshaking.

It is important to recognize that communication can be culturally loaded. Understanding communication differences across cultures and learning to adjust one's communication style to minimize misinterpretation is the key. Effective cross-cultural communication consists of individuals being willing to learn and understand cultural differences.

6. DIVERSITY IN EDUCATION

Schools are reflective of the larger society in its diversity. The changing face of society and schools was predicted as early as the 1980s. Much of the past 20 years has focused on how to best serve an increasingly diverse population in the schools. An increasing achievement gap between White students and students of color continues to be problematic. Overrepresentation of certain ethnic minority groups in special education categories

continues despite efforts to remedy the situation with a requirement for nondiscriminatory assessment as part of a multifactored evaluation for special education placement. Overrepresentation of African American males for suspensions and expulsions has been well documented. Students of color are overwhelmingly tracked in lower level classes, are underrepresented in gifted and talented classes, have lower expectations placed on them by educational personnel, and have access to fewer resources than do their more affluent peers. Finally, the limited English-proficient (LEP) population is increasing. All of these issues, as well as others, make it imperative that clinicians have a solid understanding of issues of diversity and their influence and impact in educational settings.

Having cross-culturally competent personnel in education is a must. Understanding the lives of children and their families enables the provision of services that will make positive changes in their academic and personal lives. It is well known that involved parents are important to children's education. Unfortunately, parents from lower socioeconomic status and minorities have a lower level of contact with the educational systems. The challenge for professionals is to find opportunities to bridge the cultural gap between themselves and parents. Lack of contact does not equate with lack of interest. Epstein has engaged in considerable research that has been informative regarding best practice in how to work with parents. Like other areas of diversity, issues around the distribution of power, resources, and knowledge may prohibit meaningful engagement of parents in the schools. Understanding the influence of the parents' experiences with schooling may also help to explain their reluctance to become involved. This is especially important with immigrant children whose parents' ideas of involvement may mirror what they experienced in their native lands and may be inconsistent with beliefs in the United States. In addition, a large percentage of parents in the inner city are young parents who never finished school, or were unsuccessful in school, and who view schools as negative places. These parents often did not have good role models for participation and may be struggling to understand the need for parent involvement in their children's schooling.

Finding effective ways in which to connect with parents and help them to become full partners in their children's education and mental health development can be far-reaching. To do that, psychologists need to understand the social and cultural differences that may exist between the home and the school. This understanding alone could prevent miscommunication

between parents and the school. One goal should be to establish two-way communication between the school and families. It could be quite enlightening to know parents' perceptions of the school and how they believe they are (or are not) welcomed. Clinicians are often trained to "cut to the chase" and attack the "problem" when, in fact, what may be most important is to build rapport and the relationship with parents. Instead of telling parents what they should be doing, clinicians could instead ask them, "How can we assist you?" Parents who feel truly invited into the school, have a sense of self-efficacy, and perceive that they play a critical role in their children's educations are more likely to be involved in their children's schooling. It is important to acknowledge parents as important contributors to the lives of their children.

As clinicians seek to close the achievement gap between Whites and ethnic minorities and to improve the overall educational experience, they must also understand and acknowledge conditions that may be barriers to learning. Poverty is one such variable that has far-reaching implications. Children in poverty (of which ethnic minorities are disproportionately represented) experience more problems with physical health, lower cognitive ability, poor school achievement, and the like. Despite the odds against these children being successful, there is evidence that many of these children actually do succeed. Research on resiliency identifies many factors, which are termed "protective," that enable children to beat the odds. It is important for clinicians to be familiar with this literature so that they are skilled at identifying and promoting those protective factors that are present in children's lives. Research has identified the following as protective factors: social skills (e.g., empathy, communication), problem-solving skills, a sense of control, a sense of purpose, caring relationships with adults, high expectations for doing well, and opportunities for meaningful participation in clubs, teams, and/or organizations. For many culturally diverse children and families who are also in poverty, having available resources that include human resources can make an incredible difference in their lives.

7. IMPLICATIONS FOR PRACTICE

Most disciplines acknowledge the importance of having their personnel become cross-culturally competent. Although nearly everyone in the United States shares the American culture at some level, there are many who also have subcultures that are very much a part of their lives. Recognizing the similarities as well as the differences helps to bring clinicians one step closer to providing effective services to a culturally different population. Although it may seem like a daunting task, especially if clinicians are in an area with many different cultural groups, there are some ways in which clinicians can increase their skills in being effective psychologists with culturally different groups. First, they should be open to new ideas and experiences. Too often, clinicians look at the world through their own lenses. They should take time to learn about other groups and attempt to understand the world from their perspective. This does not mean that clinicians need to agree with them so long as they at least understand these other groups' perspectives. Second, clinicians should find a cultural mediator who can guide their understanding about the beliefs and practices of a particular culture with which they may be unfamiliar. This is often done through open discussions and interpersonal sharing. Clinicians should always remember that even though they do not represent everyone from that culture, members of that culture can provide them with some insight. Third, clinicians should see diversity as something to be celebrated rather than as something that is a deficit. Fourth, clinicians should view their learning experience as a journey. There is so much to know. They will not "get it" in one course. Instead, they will continually explore, learn, and explore some more. Fifth, clinicians should explore their own cultural or ethnic identities and examine their socialization with respect to majority and minority people. Until they can understand the impact of their culture and socialization on their attitudes, beliefs, and behaviors, clinicians cannot fully appreciate the cultures of others. Finally, clinicians should always remember to strive to engage in culturally responsive practices that will ultimately benefit clients in a positive way.

The United States is a pluralistic country that is rich in diversity. This is not a phenomenon but rather a reality. Understanding the cultural differences in this country, and how they may influence behavior and aspects of society, is critical. Clinicians should embrace diversity and view it as an opportunity rather than as a challenge.

See Also the Following Articles

Acculturation ■ Cross-Cultural Psychology, Overview ■ Health Psychology, Cross-Cultural ■ Prejudice and Discrimination ■ Racial and Ethnic Minorities, Counseling of ■ Training, Cross-Cultural ■ Translation

Further Reading

Banks, J. A. (2002). *Teaching strategies for ethnic studies* (7th ed.). Boston: Allyn & Bacon.

Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom*. New York: New Press.

Friere, P. (1970). *Pedagogy of the oppressed*. New York: Seabury Press.

Hooks, B. (1994). *Teaching to transgress: Education as the practice of freedom*. New York: Routledge.

Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: HarperCollins.

Ladson-Billings, G. (2001). *Crossing over to Canaan*. San Francisco: Jossey–Bass.

McIntosh, P. (1989). *White privilege and male privilege: A personal account of coming to see correspondences through work in women's studies*. Working Paper 189, Center for Research on Women, Wellesley College.

Nieto, S. (1999). *Affirming diversity: The sociopolitical context of multicultural education* (3rd ed.). New York: Longman.

Tatum, B. D. (2003). *Why are all the Black kids sitting together in the cafeteria?* New York: Basic Books.

Downsizing and Outplacement

Wayne F. Cascio

University of Colorado, Denver, Colorado, USA

GLOSSARY

burnout A gradual process of loss during which the mismatch between the needs of the person and the demands of the job grows ever greater.

distributive justice The fairness of the outcomes of decisions, for example, in the allocation of bonuses or merit pay or in making decisions about who goes and who stays in a layoff situation.

downsizing Planned changes in a firm's organizational structure that affect its use of people; in a broader sense, downsizing refers to reductions in a firm's use of assets (e.g., financial, physical, human, information), so when referring to reductions in the numbers of people in an organization, it is appropriate to use the term "employment downsizing."

knowledge-based organizations Collections of networks in which interrelationships among individuals (i.e., social networks) generate learning and knowledge.

organizational citizenship behaviors Discretionary behaviors performed outside of one's formal role that help other employees to perform their jobs or that show support for and conscientiousness toward the organization.

organizational decline An environmental or organizational phenomenon that occurs involuntarily and results in erosion of an organization's resource base.

organizational support The extent to which an organization values employees' general contributions and cares for their well-being.

outplacement As an extension of the termination process that typically includes two elements: (a) counseling employees who have lost their jobs for emotional stress resulting from the trauma of termination, and (b) assistance with job search.

procedural justice The fairness of the procedures used to make decisions; procedures are fair to the extent that they are consistent across persons and over time, free from bias, based on accurate information, correctable, and based on prevailing moral and ethical standards.

Employment downsizing is synonymous with layoffs, but it differs from organizational decline. Its extent and scope are global, and it is an ongoing feature of many organizations. Downsizing has led to a redefinition of the concept of loyalty and has led organizations to think in terms of 3- to 5-year relationships with employees. Evidence indicates that downsizing does not lead to better long-term performance of organizations, but that it may have negative effects on victims and survivors, particularly if they perceive procedural injustice in the process. Outplacement, which provides counseling for those who have lost jobs as well as assistance in job search, can help to speed the reemployment and readjustment process.

1. DOWNSIZING DEFINED

In everyday conversation, the term "downsizing" is often used as a synonym for "layoffs" of employees from their jobs. Downsizing is commonly the result of a broader process of organizational restructuring that refers to planned changes in a firm's organizational structure that affect its use of people. Such restructuring often results in workforce reductions that may be accomplished through mechanisms such as attrition, early retirements, voluntary severance agreements, and layoffs. Layoffs are a form of downsizing, but it is important to note that downsizing is a broad term that may include any number of combinations of reductions in a firm's use of assets—financial, physical, human, and information assets. Therefore, when referring to reductions in the numbers of people in an organization, it is appropriate to use the term "employment downsizing."

1.1. Downsizing and Organizational Decline

Employment downsizing is not the same thing as organizational decline. Downsizing is an intentional, proactive management strategy, whereas decline is an environmental or organizational phenomenon that occurs involuntarily and results in erosion of an organization's resource base. For example, the advent of digital photography, disposable cameras, and other imaging products signaled a steep decline in the demand for the kind of instant photographic cameras and films that Polaroid had pioneered during the 1940s. On October 12, 2001, Polaroid was forced to declare bankruptcy.

1.2. Extent and Scope of Employment Downsizing

According to the U.S. Department of Labor, approximately 2 million people have been affected by employment downsizing each year since 1996. In 2001, for example, companies announced nearly 1 million job cuts during the 3 months after the terrorist attacks of September 11. Many firms conduct multiple rounds of employment downsizing during the same year as ongoing staff reductions become etched into the corporate culture. On average, two-thirds of firms that lay off employees during a given year do so again the following year.

The phenomenon of layoffs is not limited to the United States. They occur often in Asia, Australia, and Europe as well. Japan's chip and electronics conglomerates shed tens of thousands of jobs during the first few years of the 21st century as the worldwide information technology slump and fierce competition from foreign rivals battered their bottom lines. In Mainland China, more than 25.5 million people were laid off from state-owned firms between 1998 and 2001. Another 20 million were expected to be laid off from traditional state-owned firms by 2006.

The incidence of layoffs varies among countries in Western Europe. Labor laws in countries such as Italy, France, Germany, and Spain make it difficult and expensive to dismiss workers. For example, in Germany, all "redundancies" must by law be negotiated in detail by a workers' council, which is a compulsory part of any big German company and often has a say in which workers can be fired. Moreover, setting the terms of severance is tricky because the law is vague and German courts often award compensation if workers claim that they received inadequate settlements. In France, layoffs are rare. Even if companies offer generous severance settlements to French workers, as both Michelin and Marks & Spencer did, the very announcement of layoffs triggers a political firestorm.

2. MERGERS, ACQUISITIONS, AND THE DEMISE OF LOYALTY

Worldwide, tens of thousands of mergers and acquisitions take place each year among both large and small companies. In general, after a buyout, the merged company eliminates staff duplications and unprofitable divisions. Employment downsizing is part of that restructuring process. Such restructuring often leads to similar effects—diminished loyalty from employees. Following takeovers, mergers, acquisitions, and downsizings, thousands of workers have discovered that years of service mean little to struggling managers or a new corporate parent. This leads to a rise in stress and a decrease in satisfaction, commitment, intention to stay, and perceptions of an organization's trustworthiness, honesty, and caring about its employees.

Companies counter that today's competitive business environment makes it difficult to protect workers. Understandably, organizations are streamlining to become more competitive by cutting labor costs and to become more flexible in their response to the demands of the marketplace. But the rising disaffection of workers at all levels has profound implications for employers. Perhaps one of the most fundamental issues to consider

is the new meaning of loyalty. Today, workers and managers are less loyal to their organizations than ever before. As Reichheld has shown, U.S. companies on average now lose half of their employees in 4 years, half of their customers in 5 years, and half of their investors in fewer than 12 months.

So, what is the new meaning of loyalty? Employees at all levels are still loyal, but instead of being blindly loyal to their organizations, they are loyal to a vision or to a mission. They are loyal to mentors, protégés, or team members. As a result, more and more organizations are now thinking in terms of 3- to 5-year employment relationships. Some are even adopting a strategy long used in the sports world, that is, fixed contracts with options for renegotiation and extension.

3. CORPORATE FINANCIAL PERFORMANCE BEFORE AND AFTER DOWNSIZING

In a series of studies that included data from 1982–1994, 1995–2000, and 1982–2000, Cascio and colleagues examined financial and employment data from companies in the Standard & Poor's (S&P) 500. The S&P 500 is one of the most widely used benchmarks of the performance of U.S. equities. It represents leading companies in leading industries and consists of 500 stocks chosen for their market size, liquidity, and industry-group representation. Cascio and colleagues' purpose was to examine the relationships between changes in employment and financial performance. Each year, they assigned companies to one of seven mutually exclusive categories based on each company's level of change in employment and its level of change in plant and equipment (assets). Then, they observed the companies' financial performance (profitability and total return on common stock) from 1 year before to 2 years after the employment change events. They examined results for firms in each category on an independent basis as well as on an industry-adjusted basis.

In Cascio and colleagues' most recent study, they observed a total of 6418 occurrences of changes in employment for S&P 500 companies over the 18-year period from 1982 to 2000. As in their earlier studies, they found no significant, consistent evidence that employment downsizing led to improved financial performance, as measured by return on assets or industry-adjusted return on assets. Downsizing strategies—either employment downsizing or asset downsizing—did not yield long-term payoffs that were significantly

larger than those generated by "stable employers," that is, those companies in which the complement of employees did not fluctuate by more than ±5%.

Thus, it was not possible for firms to "save" or "shrink" their way to prosperity. Rather, it was only by growing their businesses that firms outperformed stable employers as well as their own industries in terms of profitability and total returns on common stock.

3.1. Is Downsizing Always Wrong?

Sometimes, downsizing is necessary. In fact, many firms have downsized and restructured successfully to improve their productivity. They have done so by using layoffs as part of a broader business plan. In the aggregate, the productivity and competitiveness of many firms have increased during recent years. However, the lesson from the research described previously is that firms cannot simply assume that layoffs are a "quick fix" that will necessarily lead to productivity improvements and increased financial performance. The fact is that layoffs alone will not fix a business strategy that is fundamentally flawed.

Employment downsizing might not necessarily generate the benefits sought by management. Managers must be very cautious in implementing a strategy that can impose such traumatic costs on employees—on those who leave as well as on those who stay. Management needs to be sure about the sources of future savings and carefully weigh those against all of the costs, including the increased costs associated with subsequent employment expansions when economic conditions improve. Table I shows direct and indirect costs of downsizing.

4. EFFECTS OF DOWNSIZING ON KNOWLEDGE-BASED ORGANIZATIONS

Knowledge-based organizations, from high-technology firms to the financial services industry, depend heavily on their employees—their stock of human capital—to innovate and grow. Knowledge-based organizations are collections of networks in which interrelationships among individuals (i.e., social networks) generate learning and knowledge. This knowledge base constitutes a firm's "memory." Because a single individual has multiple relationships in such an organization, indiscriminate, nonselective downsizing has the potential to inflict considerable damage on the learning and

Direct and Indirect Costs of Employment Downsizing

Direct costs	Indirect costs
Severance pay in lieu of notice	Recruiting and employment costs of new hires
Accrued vacation and sick pay	Low morale; risk-averse survivors
Supplemental unemployment benefits	Increase in unemployment tax rate
Outplacement	Lack of staff when economy rebounds; training and retraining
Pension and benefit payouts	Potential lawsuits from aggrieved employees
Administrative processing costs	Heightened insecurity; reduced productivity
Costs of rehiring former employees	Loss of institutional memory and trust in management

memory capacity of the organization. Empirical evidence indicates that the damage is far greater than might be implied by a simple tally of individuals.

When one considers the multiple relationships generated by one individual, it is clear that downsizing that involves significant reductions in employees creates the loss of significant "chunks" of organizational memory. Such a loss damages ongoing processes and operations, forfeits current contacts, and may lead to forgone business opportunities. Organizations that are at greatest risk include those that operate in rapidly evolving industries, such as biotechnology, pharmaceuticals, and software, where survival depends on a firm's ability to innovate constantly.

5. THE PSYCHOLOGICAL IMPACT OF DOWNSIZING ON VICTIMS AND SURVIVORS

Downsizing exacts a devastating toll on workers and communities. Lives are shattered, people become bitter and angry, and the added emotional and financial pressure can create family problems. "Survivors," or workers who remain on the job, can be left without loyalty or motivation. A common experience is that employee morale is the first casualty in downsizing. Workplaces tend to be more stressful, political, and cutthroat than before the downsizing. Local economies

and services (e.g., human services agencies, charitable organizations) become strained under the impact on the community. Survivors often experience "burnout," that is, a gradual process of loss during which the mismatch between the needs of the employees and the demands of the job grows ever greater. Research indicates that each person expresses burnout in a unique way but that the basic themes are the same:

• *An erosion of engagement with the job.* What started out as important, meaningful, and fascinating work becomes unpleasant, unfulfilling, and meaningless.
• *An erosion of emotions.* The positive feelings of enthusiasm, dedication, security, and enjoyment fade away and are replaced by anger, anxiety, and depression.
• *A problem of fit between the person and the job.* Individuals see this imbalance as a personal crisis, but it is really the workplace that is in trouble.

Violence and sabotage are possible among victims of the downsizing. In fact, the most common precipitator of workplace violence is a layoff or firing. Sabotage, especially cyber-sabotage, has emerged as a new threat among disgruntled ex-employees. Recently terminated workers have posted a company's payroll on its intranet, planted data-destroying bugs, and handed over valuable intellectual property to competitors. Although exact numbers are hard to ascertain, computer security experts say that this is fast becoming the top technical concern at many companies.

5.1. Roles of Procedural and Distributive Justice in Employment Downsizing

Procedural justice focuses on the fairness of the procedures used to make decisions. Procedures are fair to the extent that they are consistent across persons and over time, free from bias, based on accurate information, correctable, and based on prevailing moral and ethical standards.

Distributive justice focuses on the fairness of the outcomes of decisions, for example, in allocating bonuses or merit pay or in making decisions about who goes and who stays in a layoff situation. In simple terms, it is the belief that all employees should "get what they deserve."

In the wake of decisions that affect employees, such as those involving pay, promotions, and layoffs, employees often ask, "Was that fair?" Judgments about the fairness or equity of procedures used to make decisions (i.e.,

procedural justice) are rooted in the perceptions of employees. Strong research evidence indicates that such perceptions lead to important consequences such as employee behavior and attitudes. As noted previously, when employees believe that they have not been treated fairly, they may retaliate in the form of theft, sabotage, or even violence. In short, the judgments of employees about procedural justice matter.

Procedurally fair treatment has been demonstrated to result in reduced stress and increased performance, job satisfaction, commitment to an organization, and trust. It also encourages organizational citizenship behaviors, that is, discretionary behaviors performed outside of one's formal role that help other employees to perform their jobs or that show support for and conscientiousness toward the organization. These include behaviors such as the following:

- Volunteering to carry out activities that are not formally a part of one's job
- Persisting with extra enthusiasm or effort when necessary to complete one's own tasks successfully
- Helping and cooperating with others
- Following organizational rules and procedures, even when they are personally inconvenient
- Endorsing, supporting, and defending organizational objectives

Procedural justice affects citizenship behaviors by influencing employees' perceptions of organizational support, that is, the extent to which the organization values employees' general contributions and cares for their well-being. In turn, this prompts employees to reciprocate with organizational citizenship behaviors. These effects have been demonstrated to occur at the level of the work group as well as at the level of the individual. In general, perceptions of procedural justice are most relevant and important to employees during times of significant organizational change. When employees experience change, their perceptions of fairness become especially potent factors that determine their attitudes and behaviors. Employment downsizing is a major change for people, whether victims or survivors. Thus, considerations of procedural justice will always be relevant.

6. OUTPLACEMENT: EASING THE CAREER TRANSITION PROCESS

Outplacement is an extension of the termination process. It typically includes two elements: (a) counseling the employee who has lost his or her job for emotional stress resulting from the trauma of termination, and (b) assisting with job search. Its premise is that the sooner the terminated employee is reemployed, the less time he or she has to become disgruntled, to file a lawsuit, and/or to cause problems for those who remain, particularly through violence or sabotage.

Employment downsizing often creates anxiety or resentment among remaining employees. Some workers might feel that the departing employee received unjust treatment. Others may doubt their own job security and resent the upheaval and additional workload caused by a coworker's departure.

Outplacement can help to address these issues. It provides firing managers with training and procedures to minimize trauma to the workers affected and disruption within the affected department. It helps survivors to see their organization as a fair and considerate employer. Among those downsized, outplacement can mitigate the damaging effects of unemployment on family life by including the terminated employee's spouse in counseling sessions. Subsequent career assessment and job-search assistance may include activities such as an interest inventory as well as "how to" advice on building a résumé and doing well in a job interview. These activities focus on helping the terminated employee to identify his or her strongest skills and career preferences, to negotiate job offers, and to select the best placement.

See Also the Following Articles

Career Counseling ■ Organizational Justice

Further Reading

Cascio, W. F. (1993). Downsizing: What do we know? What have we learned? *Academy of Management Executive*, 7(1), 95–104.

Cascio, W. F. (2002a). *Responsible restructuring: Creative and profitable alternatives to layoffs.* San Francisco: Berrett–Kohler.

Cascio, W. F. (2002b). Strategies for responsible restructuring. *Academy of Management Executive*, 16(3), 80–91.

Cascio, W. F., & Young, C. E. (2003). Financial consequences of employment-change decisions in major U.S. corporations: 1982–2000. In K. P. De Meuse, & M. L. Marks (Eds.), *Resizing the organization: Managing layoffs, divestitures, and closings—Maximizing gain while minimizing pain* (pp. 131–156). San Francisco: Jossey–Bass.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86,* 425–445.

De Meuse, K. P., & Marks, M. L. (Eds.). (2003). *Resizing the organization: Managing layoffs, divestitures, and closings—Maximizing gain while minimizing pain.* San Francisco: Jossey–Bass.

Greenberg, J. (1997). *The quest for justice on the job.* Thousand Oaks, CA: Sage.

Gutknecht, J. E., & Keys, J. B. (1993). Mergers, acquisitions, and takeovers: Maintaining morale of survivors and protecting employees. *Academy of Management Executive,* 7(3), 26–36.

Kanovsky, M. (2000). Understanding procedural justice and its impact on business organizations. *Journal of Management, 26,* 489–511.

Kleinfeld, N. R. (1996, March 4). The company as family no more. *The New York Times,* pp. A1, A8–A11.

Maslach, C., & Leiter, M. P. (1997). *The truth about burnout: How organizations cause personal stress and what to do about it.* San Francisco: Jossey–Bass.

Reichheld, F. F. (1996). *The loyalty effect.* Boston: Harvard Business School Press.

Schweiger, D. M., & Denisi, A. S. (1991). Communication with employees following a merger: A longitudinal field experiment. *Academy of Management Journal, 34,* 110–135.

Driving, Risky

David W. Eby

University of Michigan, Ann Arbor, Michigan, USA

GLOSSARY

countermeasures Programs or other activities designed to solve or alleviate a problem.

graduated licensing A legislative countermeasure designed to reduce risky driving and crashes among young, beginning drivers.

mobile phones Wireless communication devices that are also known as cellular telephones or cell phones. A mobile phone is either hand-held, requiring the driver to handle the phone while engaging in a conversation, or hands-free, allowing the driver to have a conversation without physically holding the phone.

occupant crash protection devices Devices that are designed to reduce the severity of injury incurred by a motor vehicle occupant during a crash. The two most common devices are safety belts and airbags.

perceived risk A driver's perception of the probability of a negative outcome for a certain driving behavior. The two most important negative outcomes for traffic safety are crashes and citations.

risky driving behaviors Actions that increase, above a particular threshold, the objective likelihood of a crash or the severity of injury should a crash occur. A driver may not consider his or her actions to be risky even though they increase his or her chances of being in a crash or becoming severely injured in a crash.

sensation seeking The behavior associated with a psychological need to seek out and engage in activities for the purpose of stimulation.

The automobile provides unprecedented personal mobility. The costs of this convenience, however, are traffic crashes and the deaths and injuries caused by these crashes. In the United States, for example, there were 6,394,000 police-reported traffic crashes, resulting in non-fatal injury to 3,189,000 people and fatal injury to 41,821 people, in the year 2000. In fact, motor vehicle crashes are the leading cause of death for people aged 4 to 33 years.

For decades, traffic safety professionals have worked to reduce the frequency of motor vehicle crashes and the severity of resulting injury without compromising personal mobility. While this problem has been approached from several perspectives, such as designing safer roads and developing occupant crash protection devices for cars, perhaps the most important viewpoint comes from behavioral scientists. Regardless of the safety features present in cars or on the roads on which they travel, cars are operated by people whose behaviors are influenced by a multitude of psychological factors. Therefore, an important component of the traffic safety problem is understanding, predicting, and modifying the behaviors of drivers.

Because of its clear influence on crashes, much research has focused on risky driving behavior. Estimates suggest that risky driving behavior causes or contributes to at least 40% of crashes. Risky driving behaviors are those actions that increase, above some threshold, the objective likelihood of a crash or the severity of injury should a crash occur. As such, a driver may not consider his or her action to be a risky one even though it increases his or her chances of being in a crash or becoming severely injured in a crash. This definition of risky driving behavior assumes a threshold from which to assess the increase in risk or crash severity. This objective threshold is set by societal standards. In the case of speeding, for example, the threshold may be the speed limit, the speed of traffic flow, or the speed that is safe for the current road or weather conditions. Thus, drunk driving, speeding, running red lights, talking on a mobile phone, and lack of safety belt use are all examples of risky driving behaviors.

1. FACTORS RELATED TO RISKY DRIVING BEHAVIOR

Certain groups of people seem to engage in risky driving more often than others. Reviewed here are four of the most common demographic and personality factors associated with risky driving.

1.1. Age

Observational and self-reported studies of driver behavior show that young people, more frequently than others, speed, travel with shorter headways, run traffic lights, use hand-held mobile phones, and fail to use safety belts. Drinking and driving is also common among young drivers, peaking between 21 and 35 years of age. These empirical studies mirror U.S. crash statistics that show that the number of motor vehicle crashes per licensed driver is highest for young drivers and declines with age, suggesting that the frequency of risky driving decreases with age.

1.2. Gender

It is well established that men, more often than women, drive after consuming alcohol, tailgate, speed, and fail to use safety belts. Interestingly, however, use of hand-held mobile phones while driving is more common among women than men, although data on the contribution of

mobile phone use to crash involvement have not yet been routinely or accurately collected. Crashes involving alcohol, speeding, and lack of safety belt use are more common for men than women. Collectively, these data show that risky driving behavior, in general, is more common among men than women.

1.3. Sensation Seeking

A large body of work has documented the fact that certain people, referred to as sensation seekers, have a psychological need for a higher level of arousal than others, leading to negative traffic safety consequences. Sensation seekers engage in new, complicated, or emotionally intense activities, such as taking risks, because of the resulting increased arousal. Thus, a sensation seeker might engage in risky driving behavior simply to experience a situation in which physiological arousal will be elevated. Sensation seeking can be measured with a test, the Sensation Seeking Scale (SSS), in which behaviors related to sensation seeking are self reported. Several studies have shown that men score higher on the total SSS than women. Studies have also documented that SSS scores tend to increase with age up to about 19 years and then decline gradually through the life span. SSS scores also correlate significantly with frequency of motor vehicle crashes, traffic citations, and many observed and self-reported risky driving behaviors such as drinking and driving, speeding, and lack of safety belt use. Thus, there appears to be a link between risky driving behaviors and the biological need to seek arousal.

1.4. Attention Deficit/Hyperactivity Disorder (ADHD)

There is some evidence that people diagnosed with ADHD may engage in risky driving behavior more often than others. Although few studies have investigated the role of ADHD in risky driving behavior, those that have reported that people with ADHD were more likely than control subjects to speed, disregard traffic control devices, drive without a license, have multiple crashes, and have multiple moving and nonmoving violations on their record. They were also more likely to have records of failure to appear in court and failure to pay fines, suggesting that people with ADHD tend to not follow through on the consequence of their violations. Because of methodological weaknesses of these studies, these results should be considered as preliminary. In

addition, the relationship between risky driving behavior and successfully treated ADHD is unknown.

2. THEORIES OF RISKY DRIVING BEHAVIOR

The last two decades have seen a profusion of theories on risky driving behavior, with no single one emerging as clearly superior. Collectively, however, these theories show various perspectives on how risk is conceptualized and its influence on driving. In addition, there are many commonalities among the theories that can help practitioners develop countermeasures for risky driving behaviors. The following sections briefly describe a sample of theories chosen to illustrate the breadth of thinking about risky driving behavior.

2.1. Zero Risk

The theory of zero risk assumes that driving behaviors result from motives, a primary one being the need to be mobile. As people drive, they experience subjectively risky situations, which over time they learn to avoid. In addition, as drivers gain experience, previously threatening situations that cannot be avoided begin to be perceived as less risky the more they encounter these situations without adverse consequences. The theory posits that people both learn to avoid risk and adapt to risky situations they encounter by a reduction in perceived risk, to the point of experiencing zero risk while driving. Risky driving behaviors, therefore, occur either because of inexperience or because of inappropriate amount of perceived risk for certain behaviors.

2.2. Risk Homeostasis

The theory of risk homeostasis maintains that people in a population have a preferred level of risk for a given activity. Rather than attempting to minimize risk while driving, people change their driving to reach the preferred level of risk. If the perceived risk of crash or citation is reduced, through roadway improvements or reductions in police presence, for example, drivers will simply increase their risky driving behaviors to compensate.

2.3. Utility Maximization

Theories of utility maximization posit that people choose driving behaviors based upon the outcome of a cost–benefit analysis, in which the cost is the perceived risk associated with certain driving behaviors. People engage in the behavior that maximizes their benefit, or utility, at the present time. Therefore, risky driving behaviors could arise from (1) assigning a low cost to the risky outcome, such as a crash, (2) perceiving a low risk for the behavior, or (3) believing that the benefits of the risky behaviors, such as saving time, outweigh the costs of the behavior.

2.4. Problem Behavior

The theory of problem behavior posits that all risky behaviors of adolescents and young adults, not just driving, are interrelated. The theory suggests that risky behaviors result from the combination of influences from three general components of an individual's life: the psychological, social, and behavioral characteristics of the individual; the social environment; and the specific situation for the risky behavior. Each of these components has influences that make risky behavior more likely and less likely. These components interact and result in a tendency to engage in or not engage in risky behaviors. An important aspect of this theory is that behavior is motivational in nature; individuals may wish to engage in risky behaviors because they fulfill certain developmental needs, such as the need for peer approval. According to the theory, risky driving behavior results from both a motivation to engage in the behavior and a lack of inhibiting influences in the personal, social, or situational contexts to prevent the behavior.

2.5. Decision Making

Decision making theories conceptualize risky driving behaviors as resulting from a decision-making process in which risk is just one of several factors that may be considered. Therefore, the cognitive abilities of people, such as the speed at which they can process information, their memory capacity, and their general knowledge of the world, influence the selection of a given behavior. Various components of the decision, including the perceived risk, are combined to determine a subjective worth for the behavior among the set of behaviors available for a given situation. In these theories, risky driving behaviors can occur for the following reasons: (1) the risky driving behavior is selected over a less-risky behavior because it affords the person greater perceived benefit, or (2) the person, relative to others, is lacking cognitive skills or abilities for good decision making.

3. COUNTERMEASURES FOR REDUCING RISKY DRIVING BEHAVIORS

In general, traffic safety countermeasures are not based on theory. However, designing psychology-based countermeasures that are effective in reducing the incidence of risky driving behavior is of great interest to traffic safety professionals. These countermeasures fall into three general classes: legislation/enforcement, licensing, and educational programs.

3.1. Legislation/Enforcement

Most traffic safety legislation is designed to prevent risky driving behavior by making it illegal to engage in these behaviors. We are all familiar with the fact that is it is illegal to drive over the speed limit, fail to stop at a stop sign, or not use a safety belt. This legislation, however, cannot be effective without enforcement. Legislation and enforcement countermeasures work on the principle of using punishment, or the threat of punishment, to reduce the likelihood of risky driving behaviors. The effectiveness of such programs has been shown in studies of drinking and driving, safety belt use, bicycle helmet use, and motorcycle helmet use. Although not commonly used, positive reinforcement has also been shown to be effective in reducing risky driving behavior. At least one study has shown the effectiveness of a variable reinforcement schedule in increasing safety belt use. In this study, police randomly pulled over drivers and rewarded them with prizes for wearing safety belts. Legislation coupled with enforcement is the most effective way to combat risky driving behavior.

3.2. Licensing

Countermeasures based on driver licensing are designed to legally restrict driving privileges until the driver meets certain criteria. Because of the high crash rate and high incidence of risky driving among young people, for example, many states have implemented a graduated licensing system. With graduated licensing, young drivers gradually gain skills, experience, and knowledge over a period of time under controlled, less risky conditions. Generally, there are three phases to graduated licensing. In the first phase, the young driver can operate a vehicle only under adult supervision. In the second phase, the young person can drive without an adult present, but his or her unsupervised driving is subject

to restrictions, such as daytime driving only. In the third phase, the driver is given full driving privileges. Usually, these programs incorporate motivation for safe driving behavior by requiring the driver to be citation- and/or crash-free for a period of time before graduating to the next phase. These programs also occasionally include educational components in which the driver attends and passes a class and/or demonstrates driving proficiency via a driving test. In summary, graduated licensing is designed to reduce risky driving among young people by rewarding safe driving behaviors and giving young drivers the opportunity to gain experience, acquire knowledge, and mature under conditions in which risky driving behavior is less likely. Evidence suggests that graduated licensing does reduce young driver crashes by as much as 25%, probably by reducing the incidence of risky driving among young people.

Suspended licensing is also used as a sanction following a conviction for certain driving violations, usually drunk driving. The removal of a person's driving privileges is intended to both separate the person from his or her vehicle so that the undesirable driving behaviors cannot occur for some period of time, and to punish the person by taking away a privilege. This sanction has been found to be modestly effective for reducing the incidence of drunk driving, both during the sanction period and after.

3.3. Educational Programs

Educational countermeasures attempt to reduce the incidence of risky driving behaviors and crashes by presenting drivers with information relevant to traffic safety. The topics of these programs vary widely. Probably the most familiar program is the driver education course that is required in most jurisdictions for beginning drivers. The effectiveness of these programs has been extensively researched and the unfortunate conclusion is that these programs do not reduce crashes or risky driving; they do, however, teach people the rules of the road and how to operate a car. Similar findings are reported for motorcycle training programs. Other programs have attempted to teach young drivers advanced skills, such as controlling a vehicle during a skid. It has been found that when compared to a control group, young drivers completing the course actually had more crashes, although the difference was not statistically meaningful. Educational programs have attempted to teach people about the negative consequences of risky driving. The most common of these programs are victim impact panels, in which a convicted drunk driver is required to listen to

statements from the family of the victim impacted by that driver's behavior. Again, evaluations of these programs have shown that they do not reduce drunk driving behavior or crashes. As a final example, another common countermeasure involves public information and education about the presence and enforcement of a particular traffic safety law. These programs are designed to inform or remind the public about specific traffic safety legislation, such as a mandatory safety belt law, and tell the public that police are making a special effort to enforce the law, that is, to create an increase in the perceived threat of enforcement. These programs have been found to be effective, when coupled with actual enforcement, in increasing safety belt use, decreasing drunk driving, and decreasing aggressive driving.

The disappointing conclusion from these examples is that, to date, few educational programs positively impact risky driving behaviors, except those coupled with legislation and enforcement. One should not, however, draw the conclusion that educational programs cannot be effective, but rather that more research is needed in order to develop effective programs. As discussed earlier, there is no unifying or even leading theory of risky driving behavior. Lacking such theory, program development is forced to proceed in an *ad hoc* fashion. Another challenge is developing programs that use cognitively appropriate language for the target group. As recently recognized by the National Highway Traffic Safety Administration (NHTSA), traffic safety programs targeted toward young people may not accommodate their cognitive abilities. Consequently, NHTSA has begun a program to apply the principles of cognitive development in the construction of young driver programs. Finally, it should be noted that measurement of the effectiveness of an educational program has several inherent difficulties. The typical outcome measures are either infrequent (crashes), requiring several years of follow-up data, or difficult to measure on the road (risky driving behaviors). Furthermore, many programs may not reach the proper audience or may suffer from a self-selection bias, compromising the validity of evaluations.

4. CONCLUSION

Advances in our understanding and prevention of risky driving behaviors lie ahead. Many of these advances will undoubtedly result from the application of psychological theory and principles. Further research is needed to develop and test theories of risky driving behaviors and continued efforts are needed to integrate these theories

into effective countermeasures. Success in these areas will greatly enhance the prevention of the death and injury toll caused by motor vehicle crashes.

See Also the Following Articles

Accidents in Transportation ■ Attention Deficit/Hyperactivity Disorders (ADHD) ■ Decision Making ■ Driving Safety ■ Transportation Systems, Overview ■ Travel Behavior and the Environment

Further Reading

Bloomquist, G. (1986). A utility maximization model of driver traffic safety behavior. *Accident Analysis & Prevention*, *18*, 371–375.

COMSIS Corporation & the John Hopkins University. (1995). *Understanding youthful risk taking and driving.* Report No. DOT-HS-808–318. Washington, DC: U.S. Department of Transportation.

Eby, D. W., & Molnar, L. J. (1999). *Matching safety strategies to youth characteristics: A literature review of cognitive development.* Report No. DOT-HS-808–927. Washington, DC: U.S. Department of Transportation.

Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth.* New York, NY: Academic Press.

Jonah, B. A. (1986). Accident risk and risk-taking behavior among young drivers. *Accident Analysis & Prevention*, *18*, 255–271.

Jonah, B. A. (1997). Sensation seeking and risky driving: A review and synthesis of the literature. *Accident Analysis & Prevention*, *29*, 651–665.

Konecni, V. J., Ebbesen, E. B., & Konecni, D. K. (1976). Decision processes and risk taking in traffic: Driver response to the onset of yellow light. *Journal of Applied Psychology*, *61*, 359–367.

Mayhew, D. R., Simpson, H. M., Williams, A. F., & Ferguson, S. A. (1998). Effectiveness and role of driver education and training in a graduated licensing system. *Journal of Public Health Policy*, *19*, 51–67.

Näätänen, R., & Summala, H. (1976). *Road-user behavior and traffic accidents.* New York, NY: Elsevier.

National Highway Traffic Safety Administration. (1995). *Understanding youthful risk taking and driving.* Report No. DOT-HS-808–318. Washington, DC: U.S. Department of Transportation.

Shope, J. T., Molnar, L. F., Elliott, M. R., & Waller, P. F. (2001). Graduated driver licensing in Michigan: Early impact on motor vehicle crashes among 16-year-old drivers. *Journal of the American Medical Association*, *286*, 1593–1598.

Simpson, H. M. (1996). Summary of key findings: Research and information needs, program and policy priorities. In H. Simpson (Ed.), *New to the road: Reducing the risk for*

young motorists (pp. 1–17). Berkeley, CA: The Regents of the University of California.

Wasielewski, P. (1984). Speed as a measure of driver risk: Observed speeds versus driver and vehicle characteristics. *Accident Analysis & Prevention*, 16, 89–104.

Wilde, G. J. S. (1982). The theory of risk homeostasis: Implications for safety and health. *Risk Analysis*, 2, 209–225.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking.* New York, NY: Cambridge University Press.

Zuckerman, M., & Neeb, M. (1980). Demographic influences in sensation seeking and expressions of sensation seeking in religion, smoking, and driving habits. *Personality and Individual Differences*, 1, 197–206.

Driving Safety

Fredrick M. Streff

University of Michigan, Ann Arbor, Michigan, USA

This article describes behavioral targets and strategies to improve driving safety. Speeding, alcohol-impaired driving, and safety belt use behaviors are discussed, as are unsafe driving behaviors among young (age 14–18 years) and older drivers (age 70+ years). For each behavior or set of behaviors, those factors that support safe driving or reduce poor driving and those that undermine safe behaviors or support unsafe driving behaviors are discussed.

GLOSSARY

alcohol-impaired driving Instances in which at least one driver involved in a crash was impaired by alcohol.

alcohol-involved crash A crash in which at least one involved driver had been drinking prior to the crash, but the level of impairment is unknown.

crash severity Measured in terms of the extent of the damage sustained by the vehicle involved in a crash and the severity of the injuries sustained by the people involved in the crash.

deterrence Describes avoidance behaviors of drivers that are formed by perceptions of enforcement levels as well as punishment likelihood and severity.

graduated licensing A licensing procedure by which young drivers are permitted driving privileges incrementally based on the successful completion of each licensing phase, allowing young drivers to experience more complex driving situations gradually.

speed-related crashes Crashes in which a driver was traveling in excess of the posted speed limit or the speed was higher than the safe driving speed for the conditions.

1. DRIVING SAFETY AS A BEHAVIORAL TARGET

Driving safety is measured objectively by the presence or absence of motor vehicle crashes and/or the severity of motor vehicle crash outcomes. Crash severity is typically measured in terms of the damage sustained by the vehicle involved in the crash and the severity of the injuries sustained by the people involved in the crash. Driving safety is also assessed by the presence or absence of precursor behaviors demonstrated to be related to crash involvement and/or severity.

According to the National Highway Traffic Safety Administration (NHTSA), in 2001 there were 6,323,000 traffic crashes in the United States resulting in nearly 40,000 deaths and an additional 2,003,000 injuries. Annually, there are approximately 20 persons killed and an additional 1585 persons injured for every 100,000 licensed drivers. The NHTSA estimates that

the total economic cost of traffic crashes in 2000 was $230.6 billion.

Although improvements in vehicle and road design have contributed much to improving driving safety, perhaps the most significant advances in driving safety in the past 15 years are behavior change programs designed to reduce the likelihood of drivers engaging in unsafe behaviors and programs designed to increase the likelihood of safe driving behaviors. Specifically, most efforts in behavior change to improve driving safety have focused on five target areas: speeding, alcohol-impaired driving, safety belt use, and driving behaviors of older drivers (age 70+ years) and young drivers (age 14–18 years).

2. SPEEDING

The NHTSA estimates that speeding (traveling in excess of the posted limit or traveling too fast for conditions) is a contributing factor in 30% of fatal crashes, resulting in approximately 12,000 lives lost in speeding-related crashes per year. Speeding reduces a driver's ability to steer around curves or objects in the roadway, extends the distance that is needed to stop a vehicle, and increases the distance a vehicle travels as a driver reacts to a situation. Moreover, the dynamic forces generated in a crash increase exponentially with the speed of the vehicle.

Speeding behaviors are motivated and reinforced by a wide array of conditions. Speeding behaviors can result from consciously accepting risks inherent in the behavior or unconsciously by misinterpreting speed cues and subsequent risks. Drivers often speed to achieve their goal of reaching a given location at a given time. Some speed because they are not attending to available speed cues. A speedometer serves as an objective and reliable cue of vehicle speed. However, subjective cues in the driving environment (e.g., the rate at which trees on the roadside pass by as one is driving) may or may not be perceived by a driver. Even if perceived, a driver may not understand the meaning of these objective and subjective speed cues, the risks associated with the cues, and what the safe driving speed may be for a given situation. Variations in risk perceptions and acceptance among drivers compound these relationships.

Expectancies play a significant role in a person's perception and acceptance of speed. For example, drivers turning into a traffic stream on a road with a 30 mph limit would not generally expect to have a vehicle approaching at 100 mph. A misperception of closing speed in this case could easily cause a traffic crash when the turning vehicle crosses into the path of the oncoming car that should have been farther away when the turn began. Another expectancy that many drivers have is that any given driver will get to his or her destination much sooner by passing traffic traveling slowly in a cue, like one would find when two lanes of a highway are merged into one because of road construction. Although there can be a measurable time savings from speeding, the gains are minimal. A vehicle traveling in an open lane at 50 mph passing a 3-mile-long line of traffic traveling at 10 mph will gain less than 2.5 min (and no doubt anger a number of people). In fact, the faster car has to pass approximately 700 vehicles to save the 2.5 min. Passing slower vehicles may create a perception of getting to the destination faster than is warranted by objective information.

The primary means applied to limit speeding behavior has been traffic law enforcement. The theoretical roots for this approach are found in learning theory. Enforcement strategies to control speeding behaviors target change in a driver's perception of the likelihood of arrest, prosecution, conviction, and the severity of the subsequent sanction. That is, the programs change the perceived reinforcing and punitive qualities of the behavioral consequences. These perceptions serve as antecedent conditions signaling a heightened probability of negative consequences resulting from speeding. That is, speed can be regulated by decreasing the perceived utility/reinforcing properties of speeding. For example, perhaps drivers would perceive fewer benefits associated with driving on an adjoining lane of slowly moving traffic if they were made aware of the small gain made. On the other hand, if drivers perceive the environment as being too dangerous to speed, they will try to avoid the negative consequences by traveling at safe, legal speeds.

When the level of enforcement activity and its concomitant publicity decrease, changes in drivers' perceptions of the risks and rewards of speeding will result in a return to the speed behaviors that were reinforced (or at least not punished) prior to the special enforcement program. These changes may occur quickly, as is the case among drivers using two-way radios to alert other drivers to the threat of enforcement. These informed drivers will slow while in the enforcement zone and resume speed once the threat has passed. These changes may also occur slowly, as is the case when enforcement patrols target an area using an intermittent enforcement schedule.

3. ALCOHOL-IMPAIRED DRIVING

According to the NHTSA, 41% of persons killed in traffic crashes in the United States in 2001 were killed in crashes involving alcohol. That is, 17,448 persons were killed in alcohol-involved crashes in 2001. This seemingly dangerous activity is motivated by the desire to travel from point A to point B and is in part controlled by the extent to which drivers perceive risk in the behavior or reward in alternate behaviors. Behavior change programs to reduce the likelihood of alcohol-impaired driving have taken the forms of reducing the response cost for alternate behaviors, programs to reduce the social acceptability of drinking and driving, interventions to reduce alcohol dependency, and enforcement campaigns to increase the perceived threat of drinking and driving behavior.

Some of the most innovative and successful programs to reduce alcohol-impaired driving have been those that reduce the response cost and even reward alternate behaviors incompatible with alcohol-impaired driving. Incentives in the form of free food and nonalcoholic beverages for persons who serve as the designated driver for a group have proven to be successful. An interesting program that prevents a person from driving while impaired is the "tipsy taxi" program. In this program, a patron who has had too much to drink is given a ride home by a volunteer. The patron or a friend then picks up the car from where it was parked at a later time.

The primary challenge to these programs is that the response cost for the desired behaviors can be perceived to be quite high. A designated driver may resent not being able to drink or become angry at boisterous drinking friends. The designated driver may feel uncomfortable to the extent that he or she can be identified and treated poorly by others in an establishment because he or she is the designated driver. In addition to the embarrassment of having to have someone else drive one home, the tipsy taxi has a significant response cost obstacle to overcome: People do not like to leave their cars on the street or in a parking lot near a bar.

Perhaps the most common programs to reduce impaired driving are those intended to persuade people that alcohol-impaired driving is an unacceptable behavior. The foci of these persuasion programs are attitudes, perceptions, and social norms associated with drinking and driving. These programs use mass media distribution of materials to educate and persuade readers to adopt the desired set of values and to engage in the desired behaviors. Although these types of information and persuasion strategies are an effective, almost necessary, adjunct to other types of programs, persuasion programs are not in themselves an effective means to change driving behavior.

As is the case for speeding behaviors, policies, laws, and their enforcement are significant components of alcohol-impaired driving reduction programs. Strengthening laws and increasing sanctions may increase their deterrent value in two ways. First, a law may serve to discourage a group of persons from engaging in the illegal act (general deterrence). Second, a law may serve to affect the behavior specifically of persons who have had a previous experience with violating the law (specific deterrence). For example, a new law that affects only repeat offenders may have a general effect (a reduction in drinking and driving among persons without a conviction) in addition to its desired effect on the target group of repeat offenders (specific deterrence).

Behavior change programs targeting a reduction in alcohol-impaired driving also include interventions designed to reduce alcohol dependency. The most significant problem faced by treatment-based programs to reduce alcohol-impaired driving is the fact that often participants are ordered by the court to attend the treatment sessions that the court prescribes. The result is a client who has little motivation to participate actively in the change program imposed by the court and is being asked to make a major life change using a strategy that may well be inappropriate for that client.

4. SAFETY BELT NONUSE

When used, lap/shoulder belts reduce the chance of death or serious injury from a traffic crash by 45–60%. The NHTSA estimates that in 2001, use of safety belts prevented more than 12,000 fatalities that would have otherwise been caused by traffic crashes. Despite the demonstrated effectiveness of safety belts for preventing death and reducing injury, a substantial proportion of people do not use safety belts. Although NHTSA observation studies have shown that safety belt use has increased considerably from its 1994 level of 58% to its 2002 level of 75%, one out of four vehicle passengers still travel without this proven life-saving device.

Programs to promote safety belt use have focused primarily on managing short-term consequences of safety belt use. The most significant consequence, loss of life, is an extremely rare event when considered on a

trip-by-trip basis. Because of its infrequency, this consequence does little to promote belt use. In fact, prior to the implementation of laws mandating safety belt use, the national rate of belt use was approximately 14%. This rate shows that other, short-term consequences are necessary to increase belt use. Perhaps the most underused programs for promoting safety belt use are incentive programs in which a person or group of persons are rewarded for belt use.

Incentive programs have proven to be effective in achieving belt use in the 90–100% range at industrial sites and within communities. At the heart of the programs is the delivery of rewards for belt use rather than punishments for belt nonuse. Some jurisdictions have used police patrols to pull vehicles out of the stream of traffic to reward drivers, similar to the manner in which drivers are given tickets. These programs often create resentment toward police because the circumstances surrounding the delivery of the reward are perceived to be more punitive than the value of the reward. Strategies in which the delivery of rewards is concurrent with another activity in which belt use can be assessed easily by observation (e.g., fast-food drive-throughs) are highly successful.

The regularity and generalization of belt use can be accomplished with contingency management. Delaying the reward for belt use until multiple examples of the behavior are observed or using an intermittent reinforcement schedule can serve to extend the effects of the program past the time the program has ended. An example of delayed reward is a program in which persons who are observed wearing their belts receive a single bingo letter each time they order from the drive-through. The completed card may be redeemed for a product from the participating store. An intermittent schedule is achieved when belt users select a card randomly from a set of cards in which there are winning and losing cards.

Using a mixed cognitive and behavioral approach provides another strategy for promoting belt use. In this type of program, persons visiting participating stores may be given an opportunity to complete a form that pledges them to use their safety belt on every trip they take. The form is divided to create a raffle ticket that is entered in a weekly drawing for nominal prizes and a rear-mirror tag displaying their commitment to belt use. The mirror tag serves two purposes. First, it serves as a cue to the driver to put on the safety belt. Second, volunteers observe belt use on streets in the area and record the plate number of every person with a mirror tag who is using a belt. This is done throughout the program period, after which a

small number of desirable prizes are awarded to persons with the highest number of observed belt use. This mixed strategy is most effective at sites where the number of observations that can be made of all participants is approximately the same (e.g., parking lots).

The effectiveness of incentive programs is often improved by using group contingencies. When group contingencies are applied, the reward is earned only when each member of the group achieves the target level of the desired behavior. In this type of program, motivation is not derived from desire to obtain the reward alone. There are additional motivations resulting from the group asserting its influence on a person who may be preventing the group from receiving its reward.

There are also a variety of belt promotion strategies that employ punishment or the threat thereof to achieve program goals. Changes in enforcement policy mandated by laws requiring belt use have resulted in the most dramatic increases in belt use among the general population. Many, if not most, who begin to use their belt after this type of law change do so because they obey laws generally. Others respond to the fact that being pulled over is unpleasant and buckle up to avoid the stress of getting a ticket. The group of nonusers consists of persons who believe it is a personal decision; those who derive satisfaction merely from "tweaking the system"; and those who, for personal reason (e.g., fear of entrapment and discomfort), choose not to use belts.

Efforts to increase belt use through the exclusive use of mass media strategies have proven to be universally ineffective. These programs are based on increasing knowledge and shaping attitudes conducive to belt use. Although ineffective when used in isolation, media strategies can be highly effective and essential components to promoting an incentive- or disincentive-based safety belt promotion program.

5. DRIVING SAFETY AMONG OLDER DRIVERS

Achieving and maintaining safe driving behaviors among older drivers is going to be a growing challenge in the future. The NHTSA reports that in the United States there were 18.9 million licensed drivers age 70 or older in 2000, a 36% percent increase since 1990. During the same time period, the general population of older persons increased 6%, whereas the number of older drivers involved in fatal crashes increased at three times this rate (18%).

The basic motivations for driving at age 70 do not differ from those of younger drivers. One drives to get from one point to another. However, driving allows older persons the mobility required to maintain their physical and mental health. Without independent transportation, older persons may feel isolated and depressed because they cannot get together with family or friends without help; their access to medical treatment and fresh foods is reduced, further increasing their risk of an early death; and often they lose feelings of maturity and value as their driving abilities restrict their driving. The privilege to drive is often an important rite of passage from youth to adulthood in the United States. In order to avoid these negative consequences, many older persons drive in situations they have difficulty managing and may not freely admit to family members or others that they have difficulty driving in these situations.

Changes in vision, reduced range of motion, and slowed reflex and reaction times are among the effects age has on the body. Cognitively, there is a reduction in attention resources required for effective perception and decision making. Moreover, as people age, the ability to divide attention between several tasks and to identify and ignore irrelevant information decreases. As drivers age, they tend to perceive a given situation as being more threatening and their ability to respond lower than is actually the case. Older persons often take a variety of medications that may affect their ability to perceive or act in a manner consistent with safe driving. These changes affect the number, cause, and severity of crashes experienced by older drivers.

Compared to younger drivers, older drivers have a lower rate of alcohol involvement in fatal crashes but are increasingly under the influence of multiple prescription drugs that may affect driving in ways that have not yet been determined. Older drivers have a crash profile similar to that of younger drivers, but the causes of these profiles are different. As drivers age, their perceptions and assessment of risk are greater than they are in reality, and their perceived skill at meeting these challenges is lower than it is objectively. Comparing crash configurations for older and younger drivers, the NHTSA has shown that vehicles with drivers age 70 or older are three times more likely to be the vehicles struck in fatal collisions and six times more likely to crash while turning left than is the case for younger drivers.

Worsening vision and flexibility caused by aging may make assessments of an oncoming vehicle's approach speed more difficult, thus requiring more time to scan the environment. This reduces the time available for decision making. Also, age-related deficits in decision-making processes may result in drivers making decisions too slowly to maneuver to avoid a collision. Finally, once the decision is made cognitively, the ability of older persons to act on that decision behaviorally is also impaired by age.

The gradual degradation of abilities necessary for safe driving creates a significant challenge. Some programs have attempted to address problems associated with age by using training sessions to increase skeletal flexibility and mental dexterity. Although these program have had limited success, it may be useful to consider a licensing system that allows drivers to exit the driving population gracefully and in a manner consistent with their condition. However, one must keep in mind that alternatives to driving oneself from point A to point B may be extremely limited. This is especially true outside large metropolitan areas. Because there are few opportunities for traveling locally for day-to-day necessities, states are reluctant to modify licensing laws that may further restrict an elderly person's ability to be independently mobile and drive.

6. DRIVING SAFETY AMONG YOUNGER DRIVERS

According to the NHTSA, motor vehicle crashes are the leading cause of death for persons age 15–20 in the United States. In 2001, 3608 drivers in this age group were killed in fatal crashes and traffic crashes injured an additional 337,000. More disturbing is the fact that although drivers age 15–20 constitute only 7% of the driving population, they are involved in 14% of all fatal crashes and approximately 20% of all reported crashes. Research has shown that teens are three or four times more likely to be involved in a crash during their first 2 or 3 years of driving compared to the general driving population.

For drivers age 15–20 involved in traffic crashes, police recorded the cause of the crash to be driver error or other factors related to the driver's behavior in 75% of cases. Compared to drivers older than age 20, young drivers are more likely to exceed posted speed limits, drive too fast for conditions, follow too closely, cross traffic lanes, pass other vehicles, and approach signals at higher speeds. Young drivers are also less likely to use their safety belts and are affected by lower levels of alcohol than are older drivers. One of the central causes of these problems is the fact that younger drivers perceive a given situation to be less

risky and their ability to handle a given situation to be greater than is actually the case.

Until recently, driver education classes were the primary means by which young drivers were trained to drive. These programs are excellent at teaching basic skills, such as accelerating, stopping, and steering, but do little to improve the necessary perceptual and decision-making skills required for safe driving. Research has shown that driver education may at best result in a small and short-lived improvement in driving safety but appears to have no effect on crash involvement in the long term. Recently, states have begun to implement a driver licensing protocol called graduated licensing.

Young drivers are involved in a higher rate of crashes because of their lack of experience driving and their heightened levels of risk taking. Graduated licensing is a licensing procedure by which young drivers are given driving privileges incrementally based on the successful completion of each licensing phase. This allows young drivers to experience more complex driving situations gradually, reducing the chance that drivers will encounter situations beyond their skill levels. Because graduated licensing practices restrict young people from driving at hazardous times, there is an additional safety benefit associated with reduced exposure.

Differences in laws from state to state make it difficult to specify each stage of graduated licensing, but the process is generally the same for all states. The process generally starts with a highly restricted learner permit that permits driving only in daylight with an adult in the car. In the second stage, drivers earn an intermediate license that limits driving without a guardian to daylight hours and still requires supervision at night. The final stage is full licensure, which occurs when a person is age 17 or older, has completed at least 6 months of the intermediate license condition, and has remained violation and crash free for the prior 3 months. Research has shown that the implementation of a graduated licensing system results in approximately 25% fewer crashes involving young drivers compared to the number of crashes before the institution of a graduated license system.

7. CONCLUSION

Safe driving requires that appropriate behaviors take place in the correct manner at the correct time. Risky driving behaviors, such as speeding, alcohol-impaired driving, and safety belt nonuse, have been proven to be best controlled by the application of learning theory models. The most successful intervention models have been those that managed relationships between the behavioral antecedents, the desired/undesired behavior, and the consequences associated with the behavior. By restricting driving to relatively safe conditions and extending driving experience over a long period of time prior to full licensure, the crash risk of young, novice drivers can be reduced significantly.

See Also the Following Articles

Accidents in Transportation ◼ Traffic Safety Assessment ◼ Transportation Systems, Overview

Further Reading

Cooper, P. (1997). The relationship between speeding behavior and crash involvement. *Journal of Safety Research*, 28(2), 83–95.

Eby, D., & Molnar, L. (1998). *Matching traffic safety strategies to youth characteristics*, NHTSA Report 808927. Washington, DC: U.S. Department of Transportation.

Evans, L. (1994). Safety belt effectiveness: The influence of crash severity and selective recruitment. In *Proceedings of the 38th annual conference of the Association for the Advancement of Automotive Medicine* (pp. 25–42). Des Plaines, IL: Association for the Advancement of Automotive Medicine.

Federal Highway Administration. (1998). *Older driver highway design handbook: Recommendations and guidelines*. Washington, DC: U.S. Department of Transportation.

Koppa, R., Holmgreen, J., & Salzmann, A. (2001). *Adaptive equipment to enhance older driver performance: A guidebook*. College Station: Texas Transportation Institute.

National Highway Traffic Safety Administration. (2001). *Achieving a high seat belt use rate: A guide for selective traffic enforcement programs*. Washington, DC: U.S. Department of Transportation.

National Highway Traffic Safety Administration. (2002). *Traffic safety facts 2001—Overview*, NHTSA Report 809476. Washington, DC: U.S. Department of Transportation.

Sadler, D., Perrine, M., & Peck, R. (1991). The long-term safety impact of a pilot alcohol abuse treatment as an alternative to license suspensions. *Accident Analysis and Prevention*, 23(4), 203–224.

Shope, J., Molnar, L., Elliott, M., & Waller, P. (2001). Graduated licensing in Michigan. Early impact on motor vehicle crashes among 16-year-old drivers. *Journal of the American Medical Association*, 286(13), 1593–1598.

Streff, F., & Eby, D. (1994). *An evaluation of the impact and effectiveness of Michigan's drunk and impaired driving laws*. Ann Arbor: University of Michigan.

Yu, J. (1994). Punishment celerity and severity: Testing a specific deterrence model on drunk driving recidivism. *Journal of Criminal Justice*, 22(4), 355–366.

Drug Abuse

Charles E. Dodgen

Caldwell, New Jersey, USA

GLOSSARY

pharmacology The study of the effects of chemical substances on living systems.

psychoactive The ability of a drug to induce subjective effects on mood and feeling.

psychosocial Reflecting the combined influences of psychological and social factors.

relapse Resumption of drug use following a period of abstinence.

withdrawal A maladaptive behavior change that occurs when blood or tissue concentrations of a drug decline in an individual.

Drug abuse, also referred to as substance or chemical abuse, is the recurrent use of a drug despite the experience of problems caused by the drug use. Difficulties arising in certain areas of a user's life are of more importance to researchers and treatment professionals than other areas for identification of a drug abuse problem. The following are types of problems that signify drug abuse: impairment meeting major responsibilities in life, such as those regarding school, work, or home; difficulties with the law and social behavior; and aggravation of physical/medical conditions due to drug use. Drug abuse is to be contrasted with drug (chemical/substance) dependence. With drug dependence, use is considered compulsive and beyond the willful control of the user. That is, someone who is drug dependent is addicted; this is thought to be a more severe condition than drug abuse. Treatment of drug abuse is accomplished primarily using a variety of counseling and psychotherapeutic techniques employed to assist the abuser to stop using the drug, to develop new behavioral and mental coping skills, and to rehabilitate his or her life from the damage caused by the substance abuse.

1. INTRODUCTION

In the discussion of drug abuse, it would be easy but inaccurate to label any regular use of a substance as abusive. Drug use in the United States is commonplace. Many people are capable of consuming drugs without developing problems. Drugs such as caffeine and alcohol, as well as prescription pharmaceutical products such as pain killing agents or antianxiety medication, are routinely and openly consumed every day in the United States (and in other countries as well). The various drugs affect the body differently and are used for specific purposes. For example, caffeine is used to remain alert and to enhance concentration, and

tranquilizers are used to quell anxiety and for relaxation. However, drugs of abuse all have in common the property that they are psychoactive. For the sake of discussion, drugs may be classified with respect to different properties; one commonly employed system is in terms of the effect of the drug on the central nervous system (CNS). The following is one such classification system, with examples of drugs in each category:

1. CNS stimulants: Cocaine, amphetamine, and caffeine
2. CNS depressants: Alcohol, barbiturates, benzodiazepines, and solvent inhalants
3. Psychotomimetics (also known as psychedelics or hallucinogens): Marijuana, LSD, and mescaline
4. Narcotics/Opioids: Opium, heroin, codeine, morphine, and methadone

Substance use typically begins in adolescence. Adolescent substance use does not appear to be random; that is, it follows a fairly predictable pattern. Adolescents tend to start using substances that are legal and widely available to adults: alcohol and tobacco. Due to the fact that these drugs are the starting point for substance use, they are referred to as "gateway drugs." In 1975, Kandel developed a stage model of progression of drug use that has since been revised:

1. Beer or wine use
2. Hard liquor or cigarette use
3. Marijuana experimentation
4. Alcohol abuse
5. Prescription drug use
6. Opiates and other illegal drugs

The vast majority of adolescents experiment with the gateway drugs at least one time. However, although most individuals try alcohol and tobacco, only for a minority of adolescents does use advance to abusive levels. As the stages advance, progressively fewer adolescents are found in each category. For example, alcohol will be tried by approximately 9 out of 10 students by their senior year in high school and cigarettes by approximately 6 out of 10 students by senior year. Opiates, at the last stage of the model, will be tried by only 1 out of 100 students by senior year.

Due to the high prevalence of substance use in the United States, it should be no surprise that substance-related problems are often encountered by mental health clinicians. The relatively high frequency with which substance-related problems are encountered by mental health professionals reflects the influence of the following factors: (i) Drug abuse has the potential to create or worsen all psychological symptoms, such as anxiety, depression, impulsive behavior, and antisocial behavior; and (ii) people seeking mental health services also tend to be at elevated risk for substance abuse problems. In other words, drug abuse harms people and contributes to psychiatric symptoms, and people experiencing psychological problems are apt to use drugs abusively.

2. RISK FACTORS AND CAUSES OF DRUG ABUSE

With any medical or mental health condition, it is desirable to determine the cause or causes of the affliction. Identifying the cause(s) helps to develop prevention strategies to limit or eliminate future cases and treatment strategies for those already affected by the condition. For example, after the discovery that an absence of insulin was responsible for type 1 diabetes, effective treatment of diabetes with externally supplied insulin became possible. In addition, research is under way to develop early identification tests for intervention strategies to prevent later development of diabetes. This research has led to the isolation of faulty antibodies believed to attack the insulin-producing cells of the pancreas. The antibodies can be detected before the person is symptomatic for diabetes; experimental treatments are being used in an attempt to prevent the development of diabetes in these high-risk individuals.

Human behavior is complex and defies easy explanation. Unlike certain physical characteristics (e.g., eye color) or physical disorders that can be traced to single genes, a disorder such as drug abuse likely represents the interaction of multiple genetic and environmental influences. Complicating things further, ethics prevents us from conducting experimental studies (involving environmental or genetic manipulation) that might help us to tease apart various possible influences. One way to attempt to identify possible causes of substance abuse is to study risk factors. Risk factors are those variables associated with increased likelihood of developing a substance use disorder. Classes of risk factors are listed here with examples in each class:

1. Peer: Peer substance use, strong attachment to peers, and positive peer attitudes about substance use
2. Parent/family: Parent substance use, positive attitude about substance use, parent tolerance of adolescent substance use, and family disruption (e.g., divorce)

3. Personal: Early (childhood) behavior problems, poor academic performance, anxiety/depression, and low self-esteem

4. Biological: Genetic predisposition to substance use (e.g., a parent is a substance abuser)

5. Community/social: Low socioeconomic status, high availability of substances, and deviant norms that encourage use of substances

Risk factors help us to understand influences to use substances, but we know many more people use them than become abusers. Therefore, the question as to who will progress beyond experimentation and casual use to the level of abuse is not answered by risk factors alone. It appears that use of substances is more a function of external risk factors, such as peer, social, and family factors; abuse of substances appears to be more a function of personal factors, such as psychiatric, behavioral, and emotional problems.

The biopsychosocial disease model is the most widely accepted model of substance abuse and addiction. It should be clear after reviewing the list of risk factors that biological, psychological, and social factors contribute to substance abuse. The biopsychosocial model is sufficiently comprehensive to include all known contributants to substance abuse.

3. COURSE OF DRUG ABUSE

Disease conditions are defined by several common factors, such as having identifiable causes, characteristic symptoms, and established treatments. In addition, diseases have an observable course. It is important to describe the course of an illness in part so that the condition can be identified (i.e., for diagnostic purposes). Also, if the untreated progression of an illness was not known, there would be no way to judge the effectiveness of treatment. Treatment interventions endeavor, essentially, to change the course of a disease. Initial attempts to describe and classify the course of alcohol abuse depicted an ever-worsening condition that eventuated in death, unless the drinking was stopped altogether. As it turns out, the long-term outcome of regular alcohol use is not certain death. Some people who use alcohol never develop problems, some who develop problems (alcohol abusers) never become addicted, and a minority of alcohol abusers (approximately one-third) exhibit the progressive deteriorative pattern of drinking. The same overall trends may be expected with other substances of abuse as with

alcohol. In 1995, Shaffer and Robbins developed a general model to describe the typical course of an addiction, consisting of the following stages:

1. Initiation: Experimentation with a drug is begun.
2. Positive consequences: At this point in the use process, only the pleasurable pharmacological and social effects of the substance are experienced.
3. Negative consequences: For those individuals who continue to regularly use the substance, eventually negative consequences are experienced in terms of health, relationships, work, school, finances, or the law.
4. Turning point: For abusers who continue despite negative consequences, there is some recognition of the damage the substance is causing in their lives and ambivalence ensues.
5. Active quitting: For some abusers, ambivalence is resolved in the direction of stopping use.
6. Relapse prevention: For those who have quit, behavior changes are maintained over time to prevent resumption of drug use.

4. ASSESSMENT OF DRUG ABUSE

In order to treat a condition, it must first be determined that a given individual has the condition; in other words, the diagnosis of drug abuse must be made. In medicine, objective tests via technologically advanced equipment (e.g., x-ray and magnetic resonance imagery) are often used to assist the doctor in the diagnostic process. In the evaluation of drug abuse, modern technology is hardly relevant. Biological testing, in the forms of urinalysis and evaluation of saliva and blood samples, may be used but are not the mainstay of assessment. Biological testing can determine if a specific drug or drug metabolite is present in a sample but cannot indicate anything about patterns of use, withdrawal symptoms, compulsive behavior, or consequences of use, all of which are important aspects to assess. Therefore, biological testing is confined to the role of confirming recent abstinence; this information is especially important in certain settings (e.g., criminal justice system and workplace) but of limited use in a drug abuse assessment. Since we are more interested in determining whether a pattern of abusive drug use is present or not, relevant information needs to be gathered. Therefore, the interview is the primary method by which information is acquired to make the diagnosis of drug abuse. Typically, the diagnostic interview is conducted with the person in question as well as with

others in a position to observe relevant behaviors (most often family members and/or close friends). In addition to the interview, information is sometimes acquired via self-report, paper-and-pencil tests. The following information is typically obtained during a drug abuse assessment:

1. List all substances ever used
2. Age of first use of all substances
3. How used each substance (e.g., smoke, drink, snort, etc.)
4. Age of peak use, and amount used, for each substance
5. Number of days use substance per week, for each substance
6. Amount of substance used on a typical day of use
7. Date of last use of each substance
8. List all negative consequences resulting from use of substances

Diagnosing a drug abuse disorder is only one element of the assessment process. It is also necessary to determine as part of the evaluation the most appropriate setting in which treatment should take place (e.g., outpatient, halfway house, or inpatient); the proper intensity of treatment (e.g., daily treatment or monthly treatment); whether other treatment needs exist (e.g., medical and/or psychological disorders); and specific, individual treatment goals for a given person.

5. TREATMENT OF DRUG ABUSE

There is no one treatment for drug abuse. This fact is a reflection of the complexity of the condition and its diverse manifestations, and it highlights the importance of the assessment process, which is critical in helping determine the best treatment for a given individual. The treatment of drug abuse may occur in different settings, with varying degrees of professional assistance (e.g., self-help/12-step and professional help) and different modalities of professional services (e.g., individual therapy, group therapy, family therapy, and pharmacological treatment). Drug abuse treatment may be characterized as specialized treatment with one main goal: to stop the use of the substance. Treatment is primarily talking therapy—counseling and psychotherapy; in addition, medications may be employed to manage detoxification from some drugs and/or to treat coexisting psychological or medical conditions. However, regardless of the setting of treatment, the intensity of the contact schedule, or who renders the treatment, it is ultimately talking therapy that takes place. Especially early in treatment, the focus of discussion is on behavior directly related to drug use and stopping the use of the drug. Most programs and professionals recommend complete abstinence from drugs; some have the goal of harm reduction (allowing use to continue while attempting to reduce drug use to less harmful levels), but they are in the minority. As treatment progresses, and abstinence is achieved and maintained, the emphasis usually broadens to other areas of the person's life that may need repair, such as their decision-making skills, coping skills, emotional state, and relationships. In other words, the individual suffers psychological and social damage from drug abuse and may even have had significant deficits in these areas prior to his or her drug abuse; treatment is designed to improve the psychosocial functioning of the individual once he or she is drug-free.

6. RELAPSE PREVENTION

Drug abuse has been described as a chronic, relapsing disorder. Like all chronic conditions, long-term effort must be applied for the individual to maintain abstinence from drug use. Nobody would expect the blood sugar levels of someone with diabetes to be in a healthy range if the person only complied with the prescribed care regimen for 1 month after a visit to the physician. Likewise, if a drug abuser only applies the principles of treatment for a limited period of time, resumption of abusive habits would be expected. One way to attempt to guard against a backslide into prior behavior is to extend treatment as long as possible. In addition, teaching relapse prevention skills that an abuser may use going forward in time is an integral part of drug abuse treatment. Some common elements of relapse prevention programs include identification of high-risk situations that are likely to lead to relapse, development and practice of skills to effectively cope with risky situations, enhancement of self-confidence to be able to apply coping skills when needed, learning to limit a slip to an isolated incident rather than allow it to be the beginning of a process of abuse, drug/alcohol monitoring for abstinence verification, and developing positive behaviors (e.g., working and physical exercise).

See Also the Following Articles

Alcohol Dependence ■ Diagnostic and Statistical Manual of Mental Disorders ■ Drug Dependence ■ Nicotine Addiction

Further Reading

Bukstein, O. (1995). *Adolescent substance abuse: Assessment, prevention and treatment.* New York: Wiley.

Dodgen, C. E., & Shea, W. M. (2000). *Substance use disorders: Assessment and treatment.* San Diego: Academic Press.

Gold, M. S. (1991). *The good news about drugs and alcohol: Curing, treating and preventing substance abuse in the new age of biopsychiatry.* New York: Villard.

Kandel, D. (1975). Stages in adolescent involvement in drug use. *Science, 190,* 912–914.

Schuckit, M. A. (1995). *Educating yourself about alcohol and drugs: A people's primer.* New York: Plenum.

Shaffer, H. J., & Robbins, M. (1995). Psychotherapy for addictive behavior: A stage-change approach to meaning making. In A. M. Washton (Ed.), *Psychotherapy and substance abuse: A practitioner's handbook* (pp. 103–123). New York: Guilford.

Drug Dependence

Douglas H. Ruben

Best Impressions International, Incorporated, Okemos, Michigan, USA

GLOSSARY

drug tolerance The physiological buildup of a drug in the body's metabolic system generating a constant urgency for more of the drug to avoid withdrawal symptoms. Drug abusers "tolerate" more of an ingested chemical as the body habituates to its effects.

drug-dependent behavioral patterns The functional and topographical sequence of responses conditioned as a by-product of repeated episodes of drug consumption. The sequence of "behavioral patterns" consists of orderly connected responses overtly and covertly occurring from the onset of cravings to the offset of pleasure gratification. Intravenous heroin users, for example, may (i) search for a syringe (or other paraphernalia), (ii) secure an isolated place for drug preparation, and (iii) go through ritualistic steps to penetrate the vein (tighten a large rubber band around the vein site; hit the vein with two fingers to draw blood flow; insert the needle inside the vein; and compress the plunger, forcing the liquid through the syringe).

drug dependence-induced disorders Medical and psychological complications result from chronic drug dependence producing a cluster of symptoms otherwise not present in the addict's life. Common disorders include mood, dementia, Korsakoff's, anxiety, depression, psychosis, and sexual.

polysubstance addiction Progression of one drug dependence to another drug dependence so that both drugs produce similar effects, opposing effects, or complementary effects. Cross-addicted users who drink alcohol (depressant) may also enjoy sedative effects of cannabis or contrast alcohol with cocaine (stimulant–euphoric); they may also find that alcohol's calming effects take the edge off of painful crack-cocaine withdrawal.

recidivism The gradual relapse of drug addiction after a sub-stantive period of abstinence. Recurrence of drug-dependent behavioral patterns starts the spiral of habits, followed by small amounts of the drug being ingested. As drug-intake levels increase to pretreatment amounts, self-control erodes into compulsive fulfillment of cravings.

remission Period of drug treatment when addictive habits and behavioral patterns disappear but are not extinct. Medical or social stressors may force rapid detoxification and drug-free behaviors atypical in the addict's normal life. Incarceration for drug possession and intent to deliver (drug sales), for example, may last several months, dra-matically eliminating drug use. Abstinence occurs, but only artificially and temporarily since the addict restores drug dependence once released from jail.

routes of administration Entry routes of drug use include major sensory receptors such as inhaling or "snorting" through the nose (smell, olfactory), swallowing through the mouth (taste, gustatory), rubbing on skin (topical, tactile), injecting needles through skin (intravenous, intramuscular, or subdermal), or forceful anal insertion (suppository).

setting events Environmental, physiological, and psychologi-cal variables mediating the relationship between drug-dependent behaviors and their consequences. Location of drug use (e.g., home and bars), people around whom use occurs (e.g., friends and relatives), time of day, light or

darkness in room (media of contact), nervousness or indigestion (e.g., irritable bowel), migraine headaches, chronic back pain, spousal arguments, and persistent sexual deprivation are all examples of the endless continuum of drug-causing antecedents.

substitutive effects Spontaneous regeneration of drug cravings or even pretreatment use levels may result from addicts either consuming what they think are benign compounds or engaging in benign activities, all of which restore the body's climax of pleasure previously evoked by the drug. Recovering cocaine addicts who increase daily numbers of cigarettes smoked or drink more cups of caffeine-loaded coffee may unknowingly reactivate the body's tolerance for stimulation similar to using cocaine.

titration Prescribed or self-induced weaning off of one drug while simultaneously ingesting a new drug. The replacement drug may generate milder or controlled effects, such as when highly regulated methadone doses replace opioids (e.g., heroin). Titration from medication dependence may involve a gradual reduction schedule of thinning the dose or altering the frequency of intake.

withdrawal syndrome Deliberate or accidental cessation of continuous drug use causing the body to crave replenishment of the drug. Substances taken in greater amounts over a long period of time and then sharply reduced or stopped continue in the bloodstream until slowly diffused; this is also called detoxification.

Drug dependence is a pathological condition marked by biological tolerance and withdrawal symptoms resulting from consuming large amounts of a particular drug through various routes of administration (e.g., oral and intramuscular). Essential features of biological tolerance include intense and compulsive cravings for substances to reach a desired peak effect. This is due to two reasons. First, attainment of effect level, after repeated use, requires higher amounts of the same drug, whereas less amounts produced pleasure during earlier stages of use. Second, pleasurable effects diminish after repeated use of a drug, fooling the user into ingesting higher doses to restore a prior peak effect. Chronic, uncontrolled use may be triggered by life stressors that impede social, financial, occupational, and interpersonal functioning, typically impacting nuclear family members. Life stressors are cues or "setting events" evoking unwanted emotional, behavioral, or cognitive states eliminated through drug use. The second defining feature of drug dependence, withdrawal syndrome, involves physiological reactions upon immediate cessation of the drug. General symptoms observed include restlessness, anxiety, insomnia, and nausea, although severity of body complications depends on the type of

drug discontinued. For example, acute withdrawal symptoms of cocaine, such as fatigue, dysphoric mood, increased appetite, and psychomotor retardation, markedly differ from cannabis withdrawal symptoms, of which there are few except for excessive sedation or appetite stimulation. Abatement of withdrawal symptoms occurs when either abusers reuse the drug in small or large doses or titrate to another drug expected to produce commensurate effects. Users under court order to "drop urine samples" (urinalysis), for example, may cease cocaine use but increase their intake of alcohol for its stimulant properties. In this way, body cravings for peak effects remain the same when satisfied by an equivalent-acting drug.

1. CONCEPT OF DRUG DEPENDENCE

Drug dependence is a clinical term denoting significant distress or impairment in cognitive, behavioral, and physiological symptoms of users consuming large amounts of legal or illegal compounds. "Impairment," conceptually, considers the abuser to persist habitual drug ingestion despite realization of the untoward physical or psychological effects produced. Adverse effects are manifold but functionally are uniform. First, abusive use increases in quantity and frequency to match the gratifying or intoxicating effects desired. Mathematically, multiple daily routines build rapid metabolic adjustment or "tolerance" to the substance. The demand for more substance is higher to compensate for the body's adjustment (tolerance) and ensure euphoric or intoxicating sensations. Tolerance develops into dependence in two ways: physically and psychologically.

A physical dependence defines when substance intake levels greatly increase to maintain a desirable, sensory peak effect and prevent withdrawal symptoms. A psychological dependence is not biochemical in origin but contains a learning cycle of achieving optimal levels of functioning while intoxicated. Stimulant properties of high alcohol consumption, for example, may enable socially heroic, gregarious, or flirtatious behaviors in people who are introverts when sober. Withdrawal effects occur due to abrupt cessation of either physical or psychological dependence. Discontinuation causes adverse cognitive, behavioral, and physiological changes due to the decline of blood and tissue concentrations and musculoskeletal constrictions and dilations.

Withdrawal symptoms for alcohol, in particular, are classified as delirium tremens (DTs) because characteristic signs of delirium (e.g., hallucinations, delusions, and autonomic hyperactivity) are present.

1.1. Dependence versus Abuse

Drug dependence differs from drug abuse in three ways. First, drug dependence consists of clearly measurable periods of tolerance and withdrawal, whereas drug abuse sets the occasion for tolerance and withdrawal. Abuse means there exist early warning signs of life impairment predictive of physical and psychological dependence. Adults arrested for driving under the influence of alcohol, for instance, often face fears that their uncontrolled drinking may result in fatal vehicular accidents. Second, drug-dependent users show marked cognitive, behavioral, and psychomotor retardation in their daily activities. Drug abusers, however, may or may not suffer apparent cognitive, behavioral, or psychomotor malfunctioning; they may mask intoxication symptoms by performing their jobs and vehicular operations adequately and interacting normally with their families. Third, the phenomenon of withdrawal syndrome applies only to drug dependence. Drug abusers are episodic and not routine users, and their consumptive rates are variable; intermittent and lesser quantity users build resistance to adverse cognitive and visceral repercussions and only suffer mild or no side effects after abrupt cessation of desired substances.

1.2. Primary Types of Drug Dependence

The *Diagnostic and Statistical Manual of Mental Disorders* (fourth edition, text revision) nosology of drug dependence cites 11 subclassifications of substances: cannabis, opioids, nicotine, phencyclidine, sedative-hypnotics, alcohol, amphetamines, caffeine, cocaine, hallucinogens, and inhalants. Ingested routes of administration vary for each substance type, as do the dependent and withdrawal symptoms. For cannabis, oral or inhalant routes of intake of the psychoactive component tetrahydrocannabinol (THC) produce sedation, mild euphoria, and altered perception of time and sensations. Effects arise as the drug attacks the nerve cells in the brain and alters their function. Physical withdrawal effects are significant, despite the popular opinion that there are none, and include irritability, anger, depressed mood, headaches, and restlessness. Similarly, psychological effects include anxiety, paranoia, hypersensitivity, and overeating.

Opioids, whether synthetics (e.g., Vicadon), semisynthetics (e.g., heroin), or synthetics with morphine-like action (e.g., codeine), are orally, inhalantly, and intramuscularly infused. Effects include analgesic, anesthetic, and suppressive changes, whereas withdrawal signs include psychomotor agitation, impaired judgment, drowsiness, and, in severe cases, respiratory depression. Nicotine is absorbed orally (smoking or "chewing") due to the rapidity of intake and high nicotine content of the substance. Stimulant and nausea-inhibitive properties of dependence counter unpleasant withdrawal effects of insomnia, irritability, anxiety, concentration erosion, agitation, and appetite suppression. Phencyclidine (PCP or angel dust) grew from medical anesthetics in the 1950s to one of today's street drugs, producing sedative and disinhibitive effects when taken orally or intravenously. Withdrawal effects are uncertain, but stages of dependence result in rage, violence, and bizarre personality defects. Sedative–hypnotics similarly alter mood, judgment, and sexual functioning. Short-acting sedative–hypnotics (e.g., diazepam) may incur weak withdrawal effects, whereas longer acting agents cause autonomic hyperactivity, hand tremors, insomnia, anxiety, and often nausea.

Alcohol dependence depresses the central nervous system (CNS). However, to many, mild sedation under low doses is perceived as excitatory, not inhibitory. At advanced or chronic stages, there is damage to the neural pathways associated with liver failure, pneumonia, hallucinations, delusions, or head trauma. One collateral disease, Korsakoff's syndrome, arises from a thiamin deficiency from injury to the thalamus. In withdrawal, sudden termination of high alcohol intake levels can be devastating, producing an acute state of depression, convulsive seizures, nausea, agitation, and DTs.

Amphetamine dependence is bioequivalent to cocaine dependence. Both activate the CNS and generate intense but temporary anxiety, paranoia, and hypermania concurrent with aggressive, violent, or psychotic behaviors. Reversal of neuralgic arousal occurs in withdrawal phases, including lassitude, depression, weight loss, nightmares, and psychomotor retardation. Similarly, caffeine—the central ingredient in coffee, tea, colas, and most over-the-counter analgesics—ubiquitously tempts physical and psychological dependence. Substance-induced effects range from autonomic arousal and appetite suppression to inexhaustibility. Heavy doses of caffeine in excess of two or three cups of coffee daily result in restlessness, anxiety, insomnia, diuresis,

rambling speech, agitation, and possibly gastrointestinal disturbance. Eliminating caffeine "cold turkey" initiates symptoms of fatigue, anxiety, depression, or even nausea.

Unlike caffeine, hallucinogens atypically develop into physical dependence, whereas users may psychologically depend on hallucinogens for euphoric and psychedelic effects; consequently, cross-tolerance frequently exists with LSD, psilocybin, and mescaline. After stopping use, the long-acting effects of hallucinogens produce residual cravings and even persistent perception disorders or "flashbacks," a symptom commonly found during overuse or intoxication. Inhalants, too, are mood-altering, psychoactive compounds. These volatile substances, which contain esters, ketones, glycols, and halogenated hydrocarbons (found, for example, in spray-can propellants and cleaner solutions), produce inhalable vapors absorbed orally or by olfaction, reaching the lungs, bloodstream, and neural sites very quickly. Opposite effects manifest during withdrawal, including irritability, sleep disturbances, tremors, and diaphoresis.

1.3. Medication Mismanagement

Drug dependence is not always a result of teen or adult street-addictive habits. Medicines prescribed by family physicians and psychiatrists for mental health illnesses, and used appropriately at first by consumers, may be mismanaged accidentally. For example, the cerebral stimulant Ritalin, given to children with attention deficit hyperactivity disorder, requires parents to be regimen conscience. Self-administering of pain killers (Darvocet and Vicodin) also requires adherence to a rigid timetable to prevent lethal intoxication. Despite copious compliance, mistakes do occur in caretaker-administered or self-administered doses, which cause physical tolerance. The five common mismedication mistakes observed in clients are overuse, underuse, erratic use, contraindicated use, and abuse. Overuse is mistakenly taking several doses of a medication or taking the medication when it is not needed. Underuse includes both failures to fill prescriptions and forgetting to take the medication. Signs of underuse are the clients taking fewer doses than instructed to make the medication last longer or discontinuing use earlier than directed. Erratic use refers to the failure to follow physician or pharmacy instructions. This includes missing doses, taking double doses, taking doses at the wrong time, and confusion regarding which drug to take at which interval. Contraindicated use

occurs due to incorrect storage of medications, using outdated drugs, or not monitoring side effects. Errors in judgment pose higher physiological risks when clients unknowingly mix psychotropics with nonpsychotropic medicaments, such as anti-inflammatories, antihypertensives, or diuretics. The last maljudgment, that of abuse, is the intentional misuse of a drug for its secondary or tertiary properties either to hurt or help oneself. Suicidal patients overusing Tylenol with codeine may exceed the recommended dose and abuse the purpose of its analgesic properties.

1.4. Routes of Administration Accelerating Addiction

Substances ingested into the body pass through three stages before they are excreted: absorption, transmission, and metabolism. All these stages enable drugs to accelerate or decelerate through the bloodstream to the site of action. The speed with which drugs reach the site of action, called the fate of drugs, depends on how drugs enter the body or the route of administration. The three routes pertinent to drug dependence are oral, inhalation (nasal), and intramuscular. Orally taken drugs dissolve in the stomach and pass into the bloodstream through the stomach or small intestine. Solid drugs (pills and capsules) diffuse slower than liquid drugs but trigger desired effects rapidly. Inhaled drugs such as (powdered) cocaine enter the bloodstream through membranes lining the nose. The magnitude of peak effect is very fast but short in duration before excretion occurs. Consequently, abusers quickly establish a high tolerance, vigorously trying to replicate euphoric sensations they had during their first inhalant experience. By far the most accelerated drug passage to the bloodstream is intramuscularly (intravenously). Despite the obvious perils of unsanitary needles, abusers prefer injection because it bypasses the filtering process through different organs, which could retard onset of effect. Instead, the stimulated site of action is powerfully immediate, and equally brief, triggering higher and repeated injection doses for sustained pleasure.

2. CONTINGENCIES OF DEPENDENCE

The genesis of drug dependence is only partly physiological. Although drug tolerance builds body

resistance, tempting larger doses for desired effects, abusers shape their ritual habits through predictable phases of environment–behavior interactions. These phases, called contingencies, describe a unique learning relationship that the drug-using person has with his or her surroundings. Early teenage recreational use of alcohol or drugs, for example, may occur alone or with friends; this is called the setting. The teenager's disposition during these occasions is known as the set. Set and setting qualify abstractly as the place and individual emotions predisposing mild, moderate, or heavy substance use.

However, in every setting lies events. Events describe more operationally the context of settings, such as who is present, what they are doing, the time of day, and whether they are influential on the abuser. Events also name behaviors endemic to the abuser, such as deprivation and satiation levels (e.g., thirsty, hungry, full, high, and fatigued), sensory status (aroused, depressed, in pain, angry, and happy), and thoughts or expectations about the setting (plans to drink, abuse drugs, and sell drugs). Connections between setting events and behaviors create a special interdependence whereby as setting events change, they covary with changes in behaviors. In this way, settings events function like stimuli triggering not just one response but many responses in a response class. For example, metaamphetamine (crystal) abusers may buy and use the drug at a nightclub and dance with friends who also use. Drug use and dancing with friends constitute the response class. However, setting events also demarcate drug use from drug nonuse behaviors. The crystal abuser may only engage in the drug when dancing but not when he or she is alone at home and not dancing.

2.1. Types of Drug Response Classes Conditioned

Differentially learning drug habits under different settings is the result of conditioning. Conditioning means that the outcome of drug use under one setting is reinforcing and under another setting is either neutral or aversive. Simply stated, crystal abusers enjoy themselves more when they are dancing at bars rather than getting high at home. At home, the effects of crystal are benign or stagnant (neutral), or withdrawal effects may cause depression (punishment). Condition outcomes vary by another property—the response class. Responses from buying the drug or alcohol to multiple consumptions of it span a series of separate

response units held together in one of two ways. Think of a constellation of stars strung together by a common link. As one star changes, it produces specific or broad effects on surrounding stars. First, response units can be concurrent. This means two or more response units occur simultaneously rather than in a linear fashion. Crystal use may accompany respiratory elevation, muscular contractions, glandular secretions, heightened sexual arousal, speech amplification, and accelerated dancing movements. Second, response units may be sequential. Here, responses occur logically in a row like a domino effect. As one response appears, it predictably evokes the serialization of several more responses coordinated in order. After using crystal, abusers know they will perform better dance steps or feel heroically uninhibitive to flirt with the opposite sex. Response classes, either sequential or concurrent, are powerful in building drug dependence when they produce outcomes or "consequences" immediately reinforcing the final (terminal) response in a response class. When terminal responses produce aversive consequences, such as when the crystal abuser vomits violently or is robbed, drug dependence may diminish only for a short time. Drug habits regenerate rapidly because the physical cravings for the drug reactivate the same response class or a variation of the response class.

2.2. Topography and Function of Drug Response Classes

Response classes include the topography and function of behaviors. These indices measure behavior, purpose, and configurations predictive of whether drug abusers will increase or decrease consumption rates. Topography describes (i) how often the person abuses drugs (frequency), (ii) the response class (simultaneous, concurrent, etc.), (iii) how long the response class occurs (duration), and (iv) how the effects of the responses on others (magnitude). Function describes the consequences produced by the response class (reinforcing, punishing, or neutral) and whether other setting event variables interact with the response class to accelerate or decelerate occurrences. For example, topography indicates that a crystal abuser goes to a bar twice weekly, stays there for 4 h, dances with five partners, and gets irritable and aggressive toward uncooperative partners later in the evening. A functional explanation of the response class indicates that the crystal abuser dances elegantly by midevening, receives countless praise from admirers including the

partners toward whom he later makes sexual advances, and usually generates business from those wanting to learn to dance from him. In other words, functionally, drug dependence for this crystal abuser recurs partly from physical tolerance and largely from conditioned reinforcement outcomes of his drug response class.

2.3. Avoidance and Escape Effects of Drug Response Classes

Reinforcement contingencies may increase a response class by enhancing the pleasure for drug abusers, as in the case of the crystal user described previously. Reinforcement contingencies may also be effective by removing displeasure (aversive consequences) for the user. Nicotine-dependent users, for example, may smoke cigarettes both to prevent or delay fatigue onset when staying up late and to eliminate anxiety produced by daily stressors. In the first case, nicotine functionally serves to avoid aversive consequences, whereas in the second case nicotine acts as a catalyst to remove, terminate, or escape from aversive consequences. Avoidance and escape patterns become viciously intertwined in the matrix of drug response classes and account for exacerbation of physical tolerance and resistance to withdrawal effects. For example, a three-pack-a-day cigarette smoker quit (escaped) a smoking-cessation program because the withdrawal effects reminded him of the chronic anxiety suffered during his traumatic, physically abusive childhood. He used to feel nervous when battered by his mother, and the same body sensations reemerged (without him being battered) during nicotine withdrawal. Another smoker sharply increased the number of cigarettes used from 10 to 20 per day to delay (avoid) confronting his spouse about divorce. In both cases, elevated nicotine consumption proportionally correlated with avoidance and escape patterns.

2.4. Susceptibility to Drug-Dependent Response Patterns

Response classes are like an immune system. Tightly connected or interdependent responses are resilient against delays of desired drug effects, withdrawal effects, or disruptions in the response class, such as when an alcoholic runs out of beer or cannot buy more because the local stores are closed. Response classes resist extinction and the drug abuser can freely adapt without suffering withdrawal effects. In contrast, weak responses that are poorly interlinked become vulnerable to deterioration, which puts the abuser at risk of suffering withdrawal effects. Two properties of behavior causing susceptibility to degenerative drug response classes are deficits and excesses. Deficits consist of underdeveloped and immature responses falling below the social norm for adults or children and forcing abusers to expend extra effort to keep up with other people. Adults paralyzed by interpersonal deficits may be unassertive, shy, introversive, and awkward in social situations. They may drink several beers or smoke several joints (cannabis) to overcome their psychosocial deficits. However, deficiencies masked by polydrug dependence remain detrimental to abusers because their response class cannot withstand even the slightest disruptions. Two missed days of not smoking marijuana or one less beer consumed than normal may damage the response class and trigger physiological withdrawal. With excesses, the abuser overdoes responses. Excesses consist of abundant and obsessive behaviors aberrantly beyond the social norm and causing repulsion for onlookers. Compulsive exercisers, competitive athletes, academic overachievers, and perfectionists arouse constant excitement in their bodies. Excesses may be intensified or deintensified by drug dependence, but in either case the response class is fragile and susceptible to penetration. The mildest interruption of excessive behaviors (e.g., brought on by delayed access to the drug) can throw the drug abuser into a panic.

2.5. Effects of Titration on Drug Response Class

One way that drug abusers with excessive and deficient response classes protract against involuntary detoxification is by a process called titration. Titration can be controlled through medical prescription or haphazardly induced by desperate abusers. Titration consists of phasing out one drug while replacing it systematically with another drug, bioequivalently calculated to avoid withdrawal effects. Heroin (opiate synthetic) addicts traditionally taper off addictive levels through a regulated dose of a physician-prescribed drug called methadone, which is also a synthetic opiate but chemically purified to ease the abuser through stages of drug reduction. Nonprescriptive practices of titration are dangerous but common. Cocaine-dependent users may consume alcohol to offset adverse side effects of "coming down from cocaine." Consequently, the

ethanol supplement builds a tolerance in the abuser's body, eventually replacing the cocaine when cocaine supply becomes too costly or is inaccessible. Alcohol, functionally, offers risk-reduction effects for the cocaine abuser, but it does not replace the desired effects of cocaine.

When titration serves to transition one desired-effect drug to another desired-effect drug, similar to replacing heroin with methadone, the result is a substitutive effect. Substitutes occur intentionally and unintentionally. Intentional substitutes describe the cannabis abuser who abstains from marijuana to avoid THC detection on job-related random urinalyses; instead, he or she begins regular intake of two or three glasses of vodka per night to remain calm. Unintentional substitute abusers may overestimate their recovery confidence once they abstain from "bad" drugs for a period of time. Sober alcoholics, for example ("dry drunks"), may innocently consume 5–10 cups of coffee (caffeine elevation) per day, oblivious to the outcome. They may spontaneously regenerate the climax of sensory arousal previously achieved through alcohol. Renewal of these accelerated sensations (through coffee) rapidly restores body tolerance levels and makes the abuser unknowingly prone to alcohol relapse. Another common substitute misused by nondrinking alcoholics is ingestion of sweets (e.g., candy and pastries). Here, the sugar compound in sweets is bioequivalent to the sugar composition of ethanol, the deadliest antagonist to the body, causing cirrhosis. Reemergence of physical dependence on sugar, the powerful ingredient in ethanol, can swiftly sabotage sobriety by tempting nondrinkers to reconsume alcohol for satisfaction of the sugar craving.

3. FAMILY REASONS FOR RECIDIVISM

Erosion of recovery from drug dependence occurs for many reasons. The causes of recidivism range from cue-induced cravings triggered by drug substitutes and drug-using settings to unsupportive family members. The latter represents an unusual paradox in substance abuse rehabilitation because, in the spirit of change, therapists assume the victimized family welcomes abstinent members. After all, strategic planning and confrontation of the drug-abusing member by key family members probably precipitated treatment. With disunity of the family threatened, abusers are told they must receive therapy, submit to medication blockers (disulfiram, CB-1, naltrexone, methadone, etc.), or, if they refuse either of these choices, be legally or morally expelled from the family system. Ironically, compliant abusers who undergo the treatment and return to family members as caring, loving, and sober are greeted with consternation and resistance. Family members are unfamiliar with the reformed abuser's strangely alien personality and become dubious that the recovery will last. Healthy nondrug members also threaten the family equilibrium; a recovering member invokes positive and negative changes on other family members, challenging routines, rituals, traditions, and habits sanctioned for years and possibly underlying the main reason why the target member began abusing drugs. Uncertainty, skepticism, and fear become destructive agents in slowly unraveling the abstainer's confidence. Consequently, families become instigators of drug relapse in four ways, as discussed in the following sections.

3.1. Family Distrust

Distrust guards against repeated emotional and even physical injury in an unforgiving family environment. Family members who are uncertain when they will witness the next mood swing, aggressive outburst, or odd behavior of the abstainer are suspicious of every unexplained reaction. Self-defensively, family members such as children, a partner, or a spouse discredit proof of sobriety and question signs of progress. This interrogatory phase controls chaos in the family by preventing ignorance and blind faith—both of which were betrayed by abusers. Family members believe that by policing the abuser, they can monitor stability and instability of behavior and emotions and be alert to problems causing relapse. Abusers, however, object to this surveillance. They regard the distrusting family as coercive, suffocating, aggressive, and deliberately undermining their therapeutic program for recovery.

3.2. Family Protectiveness

Overtrusting may also produce untoward effects during recovery. As cohesion grows by abstainer and family members mending their wounds, the family supply of compassion goes into overdrive. Members feel compelled to help the abstainer reach recovery goals and be a productive contributor within the family. The intensity of help doubles when members fear drug recidivism and its intolerable effects on the family. They may spare the abstainer of major decision

making, financial responsibilities, and household chores and compare the drug recovery period to recuperation from surgery. Just as stitches may reopen, so it is that emotional stitches sewn together by a loose thread are perceived as fragile. The family caters to the abstainer's needs and places the recovery process on a pedestal. Caretaking, the term frequently applied to this altruism, mistakenly signals love and respect for the abstainer. The drawback of caretaking involves treating the abstainer as inactive. However, the abstainers must be active participants in altering family interactions using techniques that replace their drug abusive or dysfunctional behaviors. When families shower abstainers with unconditional nurturance, preventing natural opportunities for change, abstainers feel defeated, ineffective, and regress to pretreatment patterns.

3.3. Newly Surfaced Resentment

Family member anger kept repressed during years of the abstainer's drug dependence may surface unexpectedly when the abstainer first returns from inpatient rehabilitation or enters outpatient treatment. Previously inhibitive children, for example, may declare their resentment as victims of the abstainer's long-standing reign of terror. Embittered spouses may confess their infidelities, express thoughts of divorce, and lambaste abstainers for ruining precious years of their lives. Catharsis explodes by unloading a Pandora's box full of hatred against the abstainer with the hope of inciting debilitating guilt. Instead of shame, abstainers perceive this emotional release as a terrorist attack aimed intentionally at detonating aggression and inducing drug relapse. Two other effects on abstainers are equally disastrous. First, abstainers disenfranchise from the accusatory family harboring anger and fear. They attach onto family surrogate friends—many who abuse drugs—for emotional refuge. Second, the inescapable onslaught of allegations and criticism can result in abstainers spiraling into depression. Learned helplessness develops from suppression of their motivation and skill efforts.

3.4. Recovery as Threat to Family

Observed changes in the abstainer's behavior signal a necessary family adjustment in ways that are uncomfortable and thus avoided. Spouses, for example, are not accustomed to communicating with a healthy, sober spouse and thus resist talking. Children who never received verbal or nonverbal affection now rebuff hugs, kisses, money, or gifts given unconditionally by abstainers. Displays of love and family involvement frighten members away from the abstainer, causing the family to aggressively resent this improvement. Resistance arises when the family must engage in behaviors about which they are ignorant, have deficits, or have incompatible behaviors.

Consider affection. Instructions of therapy may inspire a wife who is a drug recoverer to be more sensual and receptive to sex. Surfacely, her romantic advances toward her husband may be what her husband always wanted. However, after years of sexual denials, emasculative teasing, and erectile inhibitions, he finds his wife's sexual overtures repulsive. He also developed alternative sexual relief outlets, such as frequent masturbation, cybersex with chat-session partners, visitation to topless bars, and flirtation at work. Consequently, his ritual of incompatible sexual behaviors is so reinforcing that the husband refuses his abstainer wife's passion and enrages her into feeling rejected and abandoned. Angered abstainers surmising their efforts are in vein may rebel against the uncooperative family members and gradually cease the recovery goals.

4. CONCLUSION

Conceptually, drug dependence is higher on the physiological and psychological continuum of pathology than drug abuse. Tolerance forms when continuous drug ingestion increases body absorption, digestion, and excretion, all directly connected to body weight, age, set and setting, gender, race, body disease, diet, personality, and history of cross-addiction. Abrupt cessation of the drug causes immediate cognitive, emotional, and biochemical changes in the body defining withdrawal symptoms. Because drug use and nonuse habits are muticausal, variables influencing high-rate or low-rate use include routes of administration and the unique relationship drug behavior shares with the environment. Under reinforcing environments, for example, drug consumption rates increase depending on the pattern or response class of drug habits. Topograhically, patterns composed of sequential or concurrent responses bond together tightly and resist extinction even when the drug abuser stops misusing the drug. Patterns of eating, sleeping, avoidance (and escape) rituals, and social routines that persist can still trigger cues inducing cravings, especially if the abuser replaces the drug with substitute compounds or titration or returns to settings that were conducive to drug use. Altering this response class thus represents the first goal of therapy, combined

with controlling the withdrawal symptoms. Moreover, conscientious therapists recognize that abuser recovery is not unilateral, wherein only the abstainer gets healthy. Instead, abuser recovery is bilateral, affecting a two-way exchange between family and abstainer. For this reason, barriers of family resistance against a reformed abuser, including distrust, protectiveness, resentment, and change, all must be addressed early in treatment to prevent hazards of drug recidivism.

See Also the Following Articles

Alcohol Dependence ■ Diagnostic and Statistical Manual of Mental Disorders ■ Drug Abuse ■ Nicotine Addiction

Further Reading

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

Ray, O. (1978). *Drugs, society, and human behavior.* St. Louis, MO: Mosby.

Ruben, D. H. (1990). *The aging and drug effects: A planning manual for medication and alcohol abuse treatment of the elderly.* Jefferson, NC: McFarland.

Ruben, D. H. (2001). *Treating adult children of alcoholics: A behavioral approach.* New York: Academic Press.

Stout, C. E., Levitt, J. L., & Ruben, D. H. (Eds.). (1992). *Handbook for assessing and treating addictive disorders.* Westport, CT: Greenwood Press.

Educational Achievement and Culture

David S. Goh

City University of New York, Queens College, Flushing, New York, USA

Harry W. Gardiner

University of Wisconsin, La Crosse, Wisconsin, USA

GLOSSARY

cultural difference theory Theory that maintains that schools in the United States are based on the majority (European American) culture, which differs in many ways from the minority cultures; these cultural differences interfere with the learning of these students and affect their motivation to do well.

deficit theory The ethnocentric idea that students who do poorly in school are considered to be deficient in some way compared with those who do well in school; most social scientists now consider this idea to be out of date.

direct participation An educational environment in which parents are actively involved in helping their children with their schoolwork, maintaining good communications with the school, participating in school activities, and the like.

ethnocentrism The belief that one's own cultural ways are superior to those of other people.

expectation theory A belief that some children do poorly in school because their teachers have a low expectation for their success.

genetic epistemology The study of the development of knowledge.

indirect support An educational situation in which there is not direct participation of the parents in their children's education; rather, parents work to maintain a home environment that is highly supportive of their children's learning (e.g., arranging for ample time to study, paying for private lessons).

individualized assessment Evaluation of individual student needs, skills, and performance.

involuntary minority groups Those who are incorporated into the U.S. society against their will (e.g., African Americans).

pedagogical model A design of systematized learning or instruction with an emphasis on the goals, principles, and methods of teaching.

standardized curricula and evaluations A set of formally designed and accepted teaching materials along with methods for measuring and comparing performance on them.

student-centered teaching Nontraditional instruction in which primary responsibility for learning is on students (rather than on teachers); students work independently or in groups on specific assignments or projects.

teacher-centered instruction Traditional approach to learning in which teachers provide information, typically through lectures along with textbooks and/or assigned readings.

triarchic model   A holistic approach to understanding differ- ences in educational achievement that looks at multiple factors (e.g., school, family, student) that contribute to ethnic differences in educational achievement.

voluntary minority groups   Those who come to the United States by their own choice (e.g., early Europeans, Asian Americans).

Culture plays a crucial role in educational achievement. Pedagogical models and practices tend to be culturally embedded. A significant portion of cross-cultural and cross-ethnic differences in achievement may be attributed to cultural factors.

1. INTRODUCTION

Educational achievement is an important concern for parents, educators, psychologists, policymakers, and others in cultures throughout the world. Many factors (e.g., pedagogical practices, educational resources, socioeconomic status [SES], parental education level) contribute to differences in the learning outcome of students. Culture has been found to play a critical role in educational achievement in cross-cultural and cross- ethnic studies. Significant differences in educational achievement have been reported among students from different countries as well as among different ethnic groups within the United States. For example, com- pared with many other countries, the United States has a wide range of ethnic groups that share within-group backgrounds, languages, values, beliefs, thinking, and norms of behavior. Available research has shown that some ethnic minority groups of students (e.g., African Americans, Hispanic Americans) achieve below the national average of all students. One exception is the high educational achievements demonstrated by Asian American students. These ethnicity-related differences have been documented in standardized achievement test scores, grades, high school graduation rates, drop- out rates, and other measures of educational achieve- ment. However, in understanding cultural or ethnic differences in educational achievement, it is important to note that there are vast variations among students within each cultural or ethnic group. Research findings that reflect between-group patterns in achievement may or may not apply to individual students within the groups. In other words, despite reported patterns of group differences, there are high-achieving students and low-achieving students within each cultural or

ethnic group. Therefore, it is important to avoid stereo- typing in understanding educational achievement and culture. One should be wary of prejudging students based on their cultural or ethnic status.

This article first presents an overview of cultural approaches to teaching, learning, and achievement. Then, it discusses research on cross-cultural differences in mathematics achievement and the explanations for the differences. This is followed by a review of different theories that have been proposed to explain cross-ethnic differences in academic achievement in the United States. In addition, the article discusses ethnic differ- ences on several cultural factors that are important to educational success. Finally, it recommends directions for future research on cross-cultural and cross-ethnic achievement.

2. OVERVIEW OF CULTURAL APPROACHES TO TEACHING, LEARNING, AND ACHIEVEMENT

North and South America represent an extremely diverse hemisphere characterized by struggles for quality and equality in education. In 2000, Mazurek and colleagues discussed educational practices and reforms in the United States, Canada, Brazil, Mexico, and Chile. They pointed out that in the United States and Canada, demands for reforms come from interest groups, the media, politicians, and parents organizations. In Brazil and Chile, national social/political movements and a return to democracy are behind educational reforms. In Mexico, reforms appear to be a response to a combi- nation of social and political turmoil within the country.

In Western Europe, education has taken on a new priority involving technology and business orientation at all levels. In some countries, this includes a partic- ular focus on vocational education to prepare students for entry into the workforce. Some countries (e.g., Switzerland, Sweden) have introduced a stronger em- phasis on vocational training in their schools, whereas others (e.g., France, Spain) continue to focus on the teaching of traditional academic subjects. At the same time, a number of Middle Eastern and African nations (e.g., Israel, Palestine, Turkey, Pakistan, South Africa), characterized by some as "nations in turmoil," are struggling to solve serious political, social, and economic problems while modernizing their countries and reforming their educational systems—a task that each country takes very seriously. At the postsecondary

level, most countries throughout the world are trying to determine what should be the proper direction and focus of their educational efforts.

2.1. Cultural Models of Schooling

Schools throughout the world are experiencing a period of rapid change and, in many cases, are finding it extremely difficult to achieve a balance among a number of critical concerns. Some of the issues that educators and schools are facing include uncertainty about what academic and cultural knowledge and skills will be needed by students in the future, wholesale revisions of curricula, experimentation in teaching strategies, the need for teachers and students to become aware of and competent in using new technologies, dramatic changes in bureaucratic and legislative policies and regulations, and increased demands on teachers.

With the exception of the education system in the United States, perhaps no other education system has been studied more intensively than that of Japan. In 2001, in a well-balanced presentation of the Japanese model of schooling, including its similarities to and differences with that in the United States, Tsuneyoshi characterized the American approach to education as one that places an emphasis on competitiveness, individual attention from teachers along with individual accomplishment on the part of students, development of cognitive abilities, and separation of teachers in terms of their disciplines. In contrast, the Japanese approach (particularly at the elementary school level) focuses on the "whole child"; close interactions between teachers and pupils for long periods of time in cooperative settings with attention to collective goals, tasks, and rewards; and efforts to provide the same or very similar treatment for all students. One advantage of the American approach that is seriously missing in the Japanese approach is the former's attention to diversity and a sensitivity and concern for minority rights. On the other hand, the Japanese approach has a clear advantage in its ability to focus on the whole child within a community structure designed to provide close, cooperative, face-to-face student–teacher relationships.

2.2. Approaches to Teacher Training

There appear to be as many approaches to teacher training as there are teacher training programs. As several recent studies show, a majority of the world's cultures have made substantive changes in teacher training and certification, whereas others are just beginning to do so. For example, in Brazil, a secondary education teacher

must pass a course of study corresponding to a teaching degree. In Russia, teachers are generally trained to teach one or two specific subjects, although many rural teachers can, and frequently do, teach between three and five subjects. An early unique feature of teacher training in Turkey was the establishment of "village institutes" where elementary teachers were trained to work with selected bright village children. Today, teachers are required to be university graduates, preferably from education faculties; however, they can receive teaching certificates if they graduate from other 4-year programs.

Yet in 1998, Murray strongly suggested that teacher education reform has largely failed, and he argued for the position that teacher education, developmental theory, and research are important and interlocking elements in bringing about a reform in school practices as well as the modification of school outcomes. It is for these reasons that Murray recommended that efforts to establish or change teacher education curricula should seriously consider the use of developmental psychology, especially pedagogical content knowledge and genetic epistemology, as a foundation.

In closing, research shows that pedagogical models tend to be embedded within cultures and that attempts to transplant them from one society to another, no matter how beneficial some may believe this will be, might not always be possible and, if attempted, will require serious planning, preparation, and attention to indigenous concerns.

Mazurek and colleagues, in discussing the debate over pedagogy, pointed out that while some countries are employing so-called progressive approaches such as student-centered teaching and individualized assessment, others are continuing to use traditional approaches such as teacher-centered instruction and standardized curricula and evaluations. Still others are experimenting with a wide variety of alternative and combined approaches. At the same time, these authors pointed to a common theme, namely that countries appear to be moving away from pedagogical strategies that are uniform and system-wide in their approach and toward those that are represented by greater diversity.

3. CROSS-CULTURAL DIFFERENCES IN MATHEMATICS ACHIEVEMENT

Research literature reports very few well-designed and carefully conducted cross-cultural studies in the area of mathematics achievement. Most notable are those carried out by Stevenson and colleagues, nearly all of which

showed American students lagging behind their counterparts in other industrialized countries. In 1990, in a classic study of math achievement among American, Chinese, and Japanese children, Stevenson and Lee attributed American children's poor performance to several factors, including a lack of emphasis on academic activities, the absence of a shared goal in academic achievement, an overestimation of children's accomplishments by parents and children themselves, low parental standards for achievement and little direct involvement by parents in their children's schoolwork, and a belief that all children, except those who may be seriously disabled, should be able to learn and do well in elementary school.

In 1998, in another study of differences in mathematical achievement among students in China, Taiwan, Japan, and the United States, Lee demonstrated that achievement differences are closely related to cultural expectations about teaching and learning. For example, in general, Asians believe and expect that everyone is capable of learning, whereas Americans believe that individual differences are inherent and, as a result, that people are expected to behave and perform differently. Lee concluded that it is this American philosophy that prevents the development of common standards and makes it difficult to institute new teaching strategies. Lee suggested that this view may seriously impair efforts to improve teaching and learning among students and teachers in the American culture for quite some time into the future. In fact, evidence from several sources shows little improvement in American mathematical performance over the past decade or so, with proficiency being even lower than it was during the 1970s.

In 2005, Gardiner and Kosmitzki reported trends similar to those just cited. In 2001, in an earlier summary of the themes and results of cross-cultural studies of mathematical achievement, Gardiner pointed first to the dominant role in this process played by mothers and second to cross-cultural evidence strongly suggesting that informal learning styles, such as those found in many Asian cultures that focusing on building interest, are more effective in teaching children a variety of skills, including math.

4. CULTURAL EXPLANATIONS FOR CROSS-ETHNIC DIFFERENCES IN ACHIEVEMENT IN THE UNITED STATES

The relationship between educational achievement and culture is also shown in research in the United States. As indicated earlier, cross-ethnic comparisons of educational achievement have shown that Asian American students outperform European American students, who in turn outperform African American and Hispanic American students. A significant portion of these differences may be attributed to cultural factors in values, beliefs, expectations, cognitive styles, and the like. These factors directly or indirectly influence students' motivation for learning and their approaches to school achievement. Over the years, several theories have been advanced to explain ethnic differences in educational achievement.

An early explanation was the so-called deficit theory. According to this theory, students who do poorly in school are considered to be deficient in some way compared with those who do well in school. The deficit may take the forms of inferior ability, culture deficiencies, and/or language problems. Proponents use these deficits to explain why some minority group students underachieve in education. These students are viewed as incapable or inferior because they exhibit cultural backgrounds, thinking, learning styles, and communication patterns that are different from those of the dominant culture. The deficit explanation stems from ethnocentrism, that is, the belief that one's own cultural ways are superior to those of other people. It is also known as the "blaming the victims" model. When children are not doing well in school, it is assumed that something is wrong with the students or with their particular culture. Most social scientists now consider the deficit explanation to be out of date.

A second explanation comes from the expectation theory, which was first made popular by Rosenthal and Jacobson in 1968. This theory posits that some children do poorly in school because their teachers have low expectations for their success. Subsequently, research has found that racial and ethnic backgrounds are factors that affect teachers' expectations of students. Teachers often have low estimations of the academic potential of students from certain racial and ethnic minority groups and do not expect much of these students. Because of the low expectations, teachers may unknowingly teach and interact with minority students in ways that are different from the ways in which they teach and interact with other students. This may, in turn, affect the students' own expectations and learning and result in low academic performance. The entire cycle becomes a self-fulfilling prophecy. The expectation theory holds that academic performance of certain minority group students can be improved if teachers can reduce their subjectivity and modify their behavior in working with these students.

The third explanation is the cultural difference theory. This theory maintains that schools in the United States are based on the majority (European American) culture, which differs in many ways from the minority cultures. Therefore, there is a mismatch or incompatibility between the school culture and the home cultures of minority students. The values, languages, cognitive styles, interaction patterns, and behaviors required in the school are often not the same as those that minority students have learned at home. These cultural differences interfere with the learning of these students and affect their motivation to do well. As a result, schools generally fail to tap the competence of minority students. Proponents of the cultural difference theory contend that it is important for teachers to develop better cross-cultural understanding and to adopt instructional practices that are more responsive to minority cultures. A learning environment that is more congruent with the minority cultures is necessary to foster effective learning of minority students and to enhance their academic achievement.

However, not all cultural differences lead to conflicts and underachievement among minority students. For example, Asian American students have cultural and linguistic backgrounds that are different from those of the majority culture in significant ways, yet these students consistently do well in school. In 1991, Ogbu proposed an explanation to account for this phenomenon. He differentiated between minority groups as involuntary immigrants and voluntary immigrants. Involuntary minority groups are those whose members are incorporated into the U.S. society against their will (e.g., African Americans), whereas voluntary minority groups are those who came to the United States by their own choice (e.g., Asian Americans). Involuntary minority students view maintaining their own cultural identities as very important and are not inclined to become enculturated to the majority culture. In addition, sometimes there is also social pressure from parents and peers discouraging these students from conforming to the school culture. For these reasons, it is not uncommon for involuntary minority students to develop a resistance to schooling and to reject behaviors that would make them successful in school. Consequently, their academic performance suffers. On the other hand, voluntary minority students have less difficulty in crossing cultural boundaries and are less affected by identity conflict. They are able to adjust to the school culture and do well in school.

Finally, a more recent explanation was proposed by Okagaki in 2001. Okagaki developed a triarchic model to explain ethnic differences in achievement. Unlike previous explanations that focus on a single perspective, this triarchic model takes a holistic approach and looks at multiple factors that contribute to ethnic differences in educational achievement. The three major factors identified in the model are characteristics of (a) the school (e.g., the structure of school and perceived function of education by various groups within a society), (b) the family (e.g., cultural beliefs and practices of parents and families), and (c) the student (e.g., ethnic and academic identities). Okagaki maintained that all three factors shape the ways in which students approach and perform in school. Therefore, all three factors need to be examined for a thorough understanding of cultural influences on minority students' achievement. Because of its comprehensive nature, the triarchic theory is able to explain some of the variations in minority students' achievement that cannot be explained satisfactorily by the earlier theories.

5. ETHNIC DIFFERENCES IN CULTURAL FACTORS AFFECTING ACHIEVEMENT

Research on ethnic differences in educational achievement over the past 20 years or so has identified a number of cultural factors that contribute to school success. Five major factors are considered in this section: educational effort, cultural values and beliefs, parental educational expectations, parental involvement, and peer influence. These factors are important because they can enhance the educational achievement of all students regardless of race or ethnicity.

5.1. Educational Effort

Educational effort has a direct impact on school performance. Students who spend more time learning and doing homework perform better than do students who spend less time studying. Educational effort has also been found to be associated with culture and ethnicity. For example, a number of studies have found that Asian American students, and to some extent European American students, are more task oriented, pay more attention in class, exert more effort, spend more time doing homework, and participate more in academic-related activities compared with other ethnic minority students. Asian Americans are also more likely to believe that high achievement is the result of motivation and hard work. On the other hand, students from other ethnic

groups are more likely to attribute academic success to ability or talent—factors that are outside of their control.

5.2. Cultural Values and Beliefs

Another factor that influences educational achievement is cultural values and beliefs related to education. All ethnic groups value education. However, in 1990, Mickelson indicated that it is not the abstract belief that education is important; rather, it is the more concrete pragmatic belief about the benefit of education that separates high-achieving students from low-achieving students. Specifically, students who do well in school believe that education serves an important function for them (e.g., getting better paying jobs, improving family financial or social conditions, overcoming occupational discrimination). Conversely, students who do not see education as critical to serving a particular pragmatic function show less motivation to do well in school. In a study conducted by Steinberg and colleagues in 1992, it was found that African American and Hispanic American students did not believe that doing poorly in school would hurt their chances for future success. Asian American and European American students, on the other hand, were more likely to believe that academic success would have a significant payoff. These different beliefs among various ethnic groups affect their behaviors related to education.

5.3. Parental Educational Expectations

Research has shown that parental educational expectations is another contributing factor to success in school and is closely associated with students' academic performance. Differences in parental expectations exist among ethnic groups. For example, compared with parents in other ethnic groups, Asian American parents place greater emphasis on educational accomplishments and have higher expectations for achievement for their children. They expect their children to receive better grades in school and to earn higher educational degrees than those of their counterparts in other ethnic groups.

5.4. Parental Involvement

Parental involvement is empirically linked to academic performance. Research has shown that parental involvement can improve children's school performance. In general, parental involvement may take place in two broad forms: direct participation and indirect support. In direct participation, parents are actively involved in helping their children with their schoolwork, maintaining good communications with the school, participating in school activities, and the like. On the other hand, indirect support does not involve direct participation of the parents in their children's education. Rather, parents work to maintain a home environment that is highly supportive of their children's learning (e.g., arranging for ample time to study, paying for private lessons). In terms of direct participation, research has shown that European American parents show the highest level of involvement in their children's education. In comparison, Asian American parents show the least direct participation in their children's education. They seem to perceive direct involvement with the school as less important due to cultural beliefs and/or language barriers. Instead, Asian American parents provide more indirect support to their children's education than do parents in other ethnic groups. Asian American parents are willing to make sacrifices for their children's educational pursuits to make sure that their children have sufficient time to do schoolwork and to create a climate in which their children's job is to study and do well in school.

5.5. Peer Influence

In addition to parental influences, studies have shown that students tend to do better with peer support for academic achievement. In some cases, peer influence is even greater than parent influence. In a 1992 study with adolescents, Steinberg and colleagues found that European American and Asian American adolescents were more likely to have friends who place a great deal of emphasis on academic achievement and, as a result, worked hard to keep up with their friends. In comparison, Hispanic and African American adolescents were less likely to receive peer support for academic achievement.

6. DIRECTIONS FOR FUTURE RESEARCH ON CROSS-CULTURAL ACHIEVEMENT

Making reliable and valid comparisons across cultures is often a bewildering and difficult task if not an impossible one. Conducting cross-cultural and cross-ethnic studies in educational achievement is fraught with a variety of methodological difficulties. Therefore, one

must exercise caution in drawing definitive conclusions from such studies. Also, as mentioned earlier, vast variation exists within any cultural or ethnic group. It is critical that research findings on achievement differences among cultural and/or ethnic groups be understood in the context of differences within groups.

Several researchers have recommended possible solutions that are worthy of further consideration. For example, in 1998, Stigler outlined the benefits and problems associated with his large-scale video survey approach to teaching and learning used in the Third International Mathematics and Science Study (TIMSS) Videotape Classroom Study. Among the advantages is the fact that video data are concrete, are grounded in practice, allow for merging of quantitative and qualitative analyses, and enable study of complex processes. However, if these data are to be used successfully, a number of disadvantages must be addressed, including sampling of only a few instructional lessons, standardization of camera procedures, problems of observer effects, sampling and validity, and confidentiality.

In another approach, in 1998, Paris and van Kraayenoord demonstrated the usefulness of specially designed tasks for assessing young children's literacy and development (e.g., asking students to examine and select books they find personally interesting and to construct meaning and stories from pictures or wordless picture books). They suggested that this approach can provide a blueprint that could lead to consistent and sensible instruction and assessment practices that are more easily understood by children, parents, and teachers.

The authors of the current article recommend that researchers seriously consider the development of a variety of new approaches, including large-scale, interdisciplinary, multimethod, longitudinal studies that focus on important aspects of educational achievement such as cooperative learning environments, multicultural literacy, and pedagogical changes in teaching and learning.

See Also the Following Articles

Academic Failure, Prevention of ■ Academic Interventions ■ Achievement Motivation in Academics ■ Educational and Child Assessment ■ Gender and Education ■ Home–School Collaboration ■ Learning ■ School–Community Partnerships

Further Reading

Alexander, R. J. (2001). *Culture and pedagogy: International comparisons in primary education*. Malden, MA: Blackwell.
Beauchamp, E. R. (Ed.). (2003). *Comparative education reader*. New York: Routledge Falmer.
Mazurek, K., Winzer, M. A., & Majorek, C. (2000). *Education in a global society: A comparative perspective*. Boston: Allyn & Bacon.
Okagaki, L. (2001). Triarchic model of minority children's school achievement. *Educational Psychologist*, *36*, 9–20.
Steinberg, L. (1996). Ethnicity and adolescent achievement. *American Educator*, *2*, 2–48.
Tsuneyoshi, R. K. (2001). *The Japanese model of schooling: Comparisons with the United States*. New York: Routledge Falmer.

Educational and Child Assessment

Jan ter Laak and Martijn de Goede

Utrecht University, Utrecht, The Netherlands

GLOSSARY

assessment A procedure whereby the assessor's reformulation of the client's problem or question is answered (or solved) using scientifically sound methods with relevant and reliable information about the child's behaviors and environment.

educational and child assessment An assessment that is focused on "maximum and typical performance," that is, on the child's intelligence, aptitude, school achievements, and typical behaviors.

individual differences approach to achievement and typical performance An approach that determines specific content domains of achievement (e.g., reading, arithmetic, biology, history, science) and of typical behaviors (e.g., socioemotional behaviors, personality); experts decide which content must be acquired as a result of a curriculum, and other experts design instruments that measure the level of acquisition of this knowledge and children's skills.

psychological assessment The professional process of investigating and judging a client's behaviors, resulting in a description, diagnosis, prediction, or explanation of the child's behaviors containing three components: implicit and explicit psychological theories and constructs, test theory representing the constructs in a quantitative model, and instruments, tests, or procedures constructed in accordance with theories and their representation in the model.

The assessment process is the professional procedure whereby a client's question (or problem) is answered (or solved) using psychological theories and constructs that are modeled by test theoretical and statistical techniques and that result in reliable and valid instruments to measure the child's achievement and typical behaviors. The goal is a valid description, categorization, prediction, and/or explanation of the child's behaviors (e.g., cognitions, emotions, observable behaviors). These are the points of departure for answering the client's question or deciding on the treatment for his or her behavioral problems.

1. INTRODUCTION

What is educational and child assessment? How do we assess the educational level of the child and his or her social behaviors and personality? In this article, first the child as the subject of an assessment is described. Second, some specific issues in assessing children are

reported, four goals of assessment are distinguished, and assessment is provisionally defined. Third, assessment is comprehensively defined with the following three components:

- Implicit and explicit behavioral theories and constructs
- Test theory and statistical models to represent these theories and constructs
- Instruments for measuring these constructs

Fourth, the assessment process, which consists of compiling these three components and contains several steps from the client's question or problem to the answer or solution, is described. This process includes the integration of diagnostic information and is subject to judgmental errors. Fifth, the possibilities and limitations of assessing achievement in children are described. Finally, the case of "Bryan" is reported using this article's definition of assessment.

Assessment in children is considered if cognitive, emotional, or behavioral problems are experienced by parents, teachers, the social environment, the institution, and/or the child himself or herself. The help of a professional assessor is sought when the implicit assessment of the parents, teachers, or child does not result in a satisfying solution. Assessment is the procedure whereby the assessor reformulates the client's question (or problem) in psychological terms and answers the question (or solves the problem) using scientifically sound methods and information about the child and his or her environment.

In assessing children, one is confronted by special issues. First, parents, teachers, and/or the school or institution—but rarely the child—ask for an answer or for help. They may or may not agree about the presence and nature of the problem. For example, Lewis reported in 1999 that the manifestation of child depression is not shared by children, parents, and teachers. Second, given that children develop rapidly and are "moving targets," how do we know whether the current behaviors are advanced or behind or whether they are normal or abnormal? An accurate picture of behavioral development is needed to make that determination. Third, assessing a child is a social process. The child must consider the assessment as a task and not as a game. Fourth, in (young) children, the nature and source of diagnostic information deviate from procedures used with adults, for example, because questionnaires cannot be used with children. Their behaviors must be observed, and parents, teachers, and peers must be interviewed.

2. FOUR GOALS OF ASSESSMENT

Although assessment is not a psychological subdiscipline such as educational and developmental psychology, it depends largely on psychology's knowledge base and methodology. Assessment refers to gathering and integrating relevant diagnostic information about the child's behaviors so as to help him or her. In their definition of developmental assessment in 1999, Johnson and Sheeber stressed the child's levels of cognitive, sensorimotor, language, and socioemotional behaviors in presenting the weak and strong sides of the child's functioning. So, assessment contains deciding, problem solving, information gathering and integrating, helping, constructing a whole picture of the child's behaviors, and using teachers, parents, peers, and the child as sources of information.

Assessment has four practical and scientific goals. The first goal is a valid description of the child's cognitions, emotions, and behaviors (hereafter referred to as "behaviors") and of his or her social environment. The practitioner needs valid descriptions that enable him or her compare children, to describe the "natural" course of development of behaviors, and to establish the effects of social contexts on behaviors. This description is also needed to estimate the relationships between behaviors, to determine context conditions that influence the behaviors, and to describe and explain their course in time. The second goal is diagnosis, which refers to how experts assign the child to a category (e.g., the attention deficit/hyperactivity disorder category). This is relevant for the practitioner and the scientist. The quality of these assignments depends on the probability of false positives and false negatives and on differential diagnoses and comorbidity of two or more behavioral syndromes. The third goal is prediction, which allows the scientist to test a nomological network and enables the practitioner to predict the child's future school achievement and social adaptation. The fourth goal is explanation, which is expressed in the hypotheses testing model of assessment. Explanation fits the dominating empirical–analytical approach in psychology.

Educational assessment is focused primarily on maximum performance. "Child assessment" in the title of this article is viewed as measuring the child's the socioemotional behavior and personality (or "typical performance"). Intelligence, aptitude, and school achievements are called "achievements."

Intelligence refers to a general ability that includes reasoning, planning, problem solving, abstract thinking,

understanding complex ideas, and learning from experience. It is not rote learning or school achievement, and it differs from creativity, character, and personality. Closely related to intelligence is aptitude, which refers to the ability to achieve in a professional field after specific training.

School achievement is related to both of these concepts but refers to what a child has achieved in a specific domain as a result of specific training or of attending a curriculum (e.g., reading or arithmetic at the primary school level, achievements at the high school or university level).

Socioemotional adaptation and personality refer to the maturity level, adaptation, lack of pathology, and individual differences in temperament and personality.

3. A COMPREHENSIVE DEFINITION OF ASSESSMENT

Psychological assessment is a process of investigation and judging that results in a description, diagnosis, prediction, or explanation of the child's (problem) "behaviors." This process involves the following three components:

- Implicit and explicit psychological theories and constructs
- Test theory representing the constructs in a quantitative model
- Test construction using theories and their representation in the model

The three components are combined in the assessment process, with the assessor acting as the "composer." The components and the process constitute the comprehensive definition of assessment used in this article and are illustrated in the following subsections. The process consists of the composition of the three components. In this process, the assessor must integrate the diagnostic information and formulate advice. Information integration is known to be subject to errors.

3.1. Psychological Theories and Constructs in Assessing Children's Achievements and Social Adaptation and Personality

Assessment involves content (i.e., behaviors) that is inferred from observations and that fits in a theory or network of constructs. The many theories and constructs

for children's achievements and personality are subsumed under the "three central orientations" or approaches used in psychology:

- The individual differences orientation
- The developmental orientation
- The (social) context orientation

The prototype of the individual differences orientation is Spearman's approach to intelligence. Spearman insisted on the study of individual differences as being the main way in which to understand intelligence. He discovered or invented the "G" (general) factor because he assumed that one general dominating ability was present in all of the activities requiring mental effort. In 1998, Jensen characterized the G factor as a discovery that can be compared to Skinner's laws of reinforcement and Thorndike's law of effect.

Thurstone was the father of the multiple intelligence theories. He accepted the G factor but attached little significance to it. He viewed it as an abstract, "second-order factor" that "floated above" more interesting specific factors. Thurstone discovered or invented the seven "primary mental abilities": verbal comprehension, word fluency, number, space, associative memory, perceptual speed, and reasoning. Later, Guilford distinguished 120 factors, and Vernon created order among the many factors by formulating a hierarchical model with G on top and two broad groups of factors below: verbal–educational (e.g., English, history) and spatial–mechanical (e.g., automotive repair). Below these, he defined all kinds of specific factors.

The orientation of Spearman and the other authors is known as the psychometric approach to intelligence. This approach combines intelligence and aptitude because Thurstone's multiple intelligence resulted in measuring aptitudes in the same way as the primary mental abilities.

The individual differences approach to school achievement does not aim to discover factors in school achievement; rather, it aims to determine specific content domains such as reading, arithmetic, biology, history, and science. Experts decide what knowledge must be acquired as a result of a curriculum, and other experts (e.g., from the Educational Testing Service [ETS]) design instruments that measure the level of acquisition of this knowledge and the children's skills. The results can be used to describe individual differences in children's achievements, the level reached by a class or school, and/or the degree of success of the curriculum and teachers. In addition, the individual differences approach has had a great impact on child

assessment in measuring dimensions of personality, temperament, and social adaptation.

This approach served the goal of providing a reliable and valid description of the ranking order of a child's achievement and personality. It helped to categorize children, for example, distinguishing cognitively retarded children from nonretarded children and distinguishing well-adapted children from not well-adapted children. The individual differences were used for predicting the future levels of achievements, school and vocational success, and levels of adaptation.

The prototype of the developmental orientation is Piaget's cognitive developmental theory. Piaget was convinced that to understand human cognition, one must study its historical and ontogenetic development (i.e., within the life course). His original question was an epistemological one: How could humankind change from biological beings to logical scientific ones? Piaget was not interested in individual differences or in fostering cognitive development by curricula or training. His subject was not a specific pupil but rather the epistemic subject, and the environmental stimuli were only an "aliment" (nourishment) to existing cognitive structures. Piaget's theory was also applied to social adaptation (e.g., the development of levels of perspective taking). This developmental orientation implied that children's stages of cognitive and social development had to be assessed. The stage concept hypothesized the existence of qualitatively different stages, ordered in time, that resulted in an inevitable and equilibrated final stage and that were characterized by sudden transitions. Empirical support for these claims was scarce and circumstantial.

This developmental orientation could be readily applied because many achievements at school (e.g., understanding simple scientific laws, understanding the concept of living beings, understanding economic laws) can be ordered along the Piagetian developmental dimension from prelogical thought to formal operational thought.

The impact of the developmental orientation on educational and child assessment is limited. Instruments for measuring stages were scarce, and measurements were more difficult to conduct and interpret than were the traitlike individual differences measurements.

The prototype of the contextual orientation to educational and child achievements is the attempt to influence achievements and socioemotional adaptation by curricula and programs. The latter are the independent variables that should change the dependent variables of intelligence, aptitude, and school achievements and

socioemotional adaptation. This orientation is also considered to be the only road to describing and explaining behaviors by efficient causes, and it fits well the dominating empirical–analytical approach in psychology. The methodology was important, because the effect of independent variables could be revealed. It served the goal of determining the effect of training, intervention, and situational change (i.e., the context) on achievements and on social adaptation and personality. These independent variables of training and the like could vary from small manipulations in content and form of curricula to big programs such as Head Start. The dependent variables were either borrowed from the individual differences orientation (e.g., intelligence, aptitude, and school achievements; level of socioemotional adaptation) or constructed for a specific question (e.g., change in a memory retrieval strategy).

The context orientation was not aimed at constructing instruments for measuring children's achievements and personality. This is possibly one reason why assessors underestimate the role of the context in understanding and explaining achievements and typical performance. However, the effects of schooling on intelligence and the effects of the cultural influence on generational growth of intelligence are present (e.g., in the Flynn effect).

Educational and child assessment is dominated by the individual differences approach, partly because this approach contained the simplest road to developing instruments for measuring maximum and typical performance in children.

3.2. Modeling Children's Achievements and Socioemotional Adaptation

3.2.1. Modeling Individual Differences in Intelligence and Aptitudes

Spearman combined conceptual analysis and test-theoretical and statistical modeling. His two-factor theory implied that individual differences in true scores could be attributed to two factors: general (present in all tasks) and specific (present in tasks of the same type, e.g., verbal, spatial). His type of factor analysis was suited to discover a first strong factor (e.g., with the now obsolete centroid method).

Thurstone's multiple intelligence construct required a search for a simple structure, that is, where items loaded high on one intended factor and low on all other factors. Orthogonal rotation and a specified number of factors

were appropriate. This simple structure concept is usually applied to the measuring of aptitudes (e.g., in the differential aptitude test). Moreover, simple structure analysis is often applied in an exploratory way to reduce many items to a few dimensions.

Guilford needed a factor analysis that discovered many factors. He "forced" the items in these factors by using the Procrustes technique. The hierarchical model was developed to rank the many factors.

The analysis of dimensionality and of the quality of individual items is now supported by the many possibilities of modern test theory or item response theory (IRT). These theories contain models to test the unidimensionality of a scale and to weight the individual items according to their contributions to the total test score.

The analysis of school achievement was a matter of experts deciding what to measure and how to measure it. There was no effort to find any dimensionality for school achievement. Modern test theory (IRT) is however, used more and more to construct one-dimensional achievement tests.

3.2.2. Modeling Development of Achievements

Developmental psychology was characterized by Cronbach as an individual differences discipline, whereas Wohlwill tried to define a real developmental orientation. By considering the developmental orientation as an individual differences orientation, developmental research was biased toward investigating stability with the help of tests, that is, rank order and normative stability. Because researchers were looking for stability, they found it. However, there was never complete stability, even after correction for measurement error. The dominating message from this research is that achievements are relatively stable, predictable, and foreseeable. Moreover, the message is that achievements increase over time in the individual and sample because children grow older and because of the Flynn effect. Rarely is a developmental construct used to model the development of achievements in the course of time.

3.2.3. Modeling the Context of Achievements

The effects of manipulations of independent variables, interventions, programs, events, stimuli, and situations on achievements can readily be modeled with (multiple) analysis of variance (ANOVA). True and quasi-experimental designs have been well elaborated. This technique is appropriate for determining the existence of significant differences at the group level, and

computing effect sizes, that give a fair estimate of gain. However, these differences do not apply to every element of the sample, and the differences do not necessarily last forever. Both are partly reasons for the disappointing results seen in programs to improve children's achievements. In addition, behaviorists' optimism for the extent to which social change can be effected by educational programs has proven to be misguided.

The modeling of social adaptation and personality does not deviate from this pattern in modeling intelligence, aptitude, and school achievements. The individual differences approach and factor analysis also rule these domains, and IRT is rarely applied in measuring personality and socioemotional adaptation.

3.3. Instruments for Assessing Children's Achievements and Socioemotional Adaptation

Theories and constructs, on the one hand, and test-theoretical and statistical modeling, on the other, can be balanced and produce theory-driven instruments for educational and child assessment. However, practical questions have dominated test construction, resulting in the gathering of items that could predict practical criteria such as school and vocational success and adaptation and pathology. Examples of instruments measuring the following eight domains are presented next:

- Screening of developmental backwardness
- Cognitive and sensorimotor development
- School achievement
- Socioemotional adaptation
- Childhood pathology
- Personality and temperament
- Social status in a peer group
- Children's social environments

3.3.1. Screening of Developmental Backwardness

Screening instruments attempt to measure developmental backwardness or disturbance. An example is the Denver Developmental Scale. Children from 6 days to 6 years 6 months of age are investigated with tasks, and the parents are interviewed. Despite its name, this instrument is an individual differences instrument and considers backwardness to be a one-dimensional scale. The McCarthy Developmental Scales presuppose

orderly development in several domains. However, this presupposed "natural" development has not been tested empirically. These instruments contain different scales for cognitive and sensorimotor behaviors, and the children (2–9 years of age) are investigated using tasks and games to elicit problem-solving behaviors. Some items must be answered by the parents. The Gesell Motor Scales measure motor, speech, and social behaviors in children from 4 weeks to 6 years of age. Gesell presupposed a developmental scale in these behaviors.

3.3.2. Cognitive and Sensorimotor Motor Development

Cognitive and motor development can be investigated in young children using the Bayley Developmental Scales. Tasks are given to babies and toddlers, and their parents are interviewed. Two scores result: one for mental development and one for motor development. This is an individual differences test, and the two factors are rationally defined. The Wechsler scales for preschool children (Wechsler Preschool and Primary Scales of Intelligence [WPPSI]) and for school children (Wechsler Intelligence Scale for Children [WISC]) are the most widely used intelligence tests. Wechsler defined intelligence as the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his or her environment. Intelligence was considered to be an individual differences phenomenon. The performance and verbal intelligence factors were rationally distinguished and, in fact, usually correlate at approximately .70. Many factor analyses were conducted. The "three Kaufman factors" (verbal comprehension, perceptual organization, and freedom from distractability) interpretation of the WISC is popular. The existence of these factors is supported by factor-analytical results. The Stanford–Binet scale is the oldest intelligence test. The intelligence quotient (IQ) is defined as a comparison between what is normal for a certain age and what is observed in a specific child.

3.3.3. School Achievement

School achievement tests (SATs) are found in every Western country and many other countries. These were developed to cover a content domain and are treated as individual differences variables. It is not the primary goal of these measures to distinguish underlying factors that generate or cause these individual differences. SATs and grade point averages are useful in selecting and placing children in various educational levels, and they do predict future school achievements fairly well. The Differential Aptitude Test contains nine subtests (e.g., sentences, analogies, technical ability, arithmetic skill, words) that resemble Thurstone's multiple intelligence factors.

3.3.4. Socioemotional Adaptation

Social emotional adaptation can be measured using many instruments, but all are based on the individual differences orientation, for example, Harter's scales for measuring the aspects of self (e.g., academic, social, and motor skills; being socially accepted).

3.3.5. Childhood Pathology

Childhood pathology can be measured by Achenbach's Child Behavior Check List (CBCL). It is based on the individual differences approach, and its factors are partly rational and empirically defined (e.g., the distinction between internalizing and externalizing behaviors). In 2000, Wenar and Kerig described several pathological syndromes in children using the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). Although this system was originally developed for persons 18 years of age or over, it can be used for childhood pathology as well. In addition, many lists are available for teachers to judge the social and task behaviors of pupils. Kovacs's questionnaire to measure depression is a downward extension of the Beck Depression Scales and allows one to define clinical cases of depression in children. The Attention Deficit/Hyperactivity Disorder (ADHD) is a short questionnaire to measure impulsivity, lack of attention, and the level of undirected activity in children 4 to 18 years of age. The Bender consists of prescribed figures that children must draw and helps to screen neurologically based deficits in perception and expression of simple and complex "Gestalts."

3.3.6. Personality and Temperament

Personality characteristics are considered to be individual differences variables. In addition to the "Big Five," there are the "Small Five" for children measuring extraversion, agreeableness, conscientiousness, emotional stability and openness. The California Q sort of Block's Berkeley group distinguished three types of children: ego resilient, overcontrolled, and undercontrolled children $2\frac{1}{2}$ to 18 years of age. These can easily be connected to the five personality factors.

Temperament is often considered to be a precursor to the personality traits. Several distinctions can be made: easy, slow to warm up, and difficult children.

3.3.7. Social Status in a Peer Group

Social status is an important variable because children have peers who influence their behaviors. In this procedure, children nominate peers with whom they like to play or work. This results in three types of children: popular, unpopular, and neglected children.

3.3.8. Children's Social Environments

Family and school relationship instruments measure individual differences in families and schools. Moos's social climate scales usually consist of approximately 100 items and result in three to nine dimensions underlying these items. The three most general dimensions are relationship (e.g., cohesion, conflict), personal growth (e.g., autonomy, task oriented), and system continuity/change (e.g., control, order, leadership). There are several questionnaires to measure the parental styles: authoritative, authoritarian, or laissez-faire. The school environment is also measured, that is, how this environment fosters achievement motivation, well-being, and self-confidence in pupils. The mental measurement yearbooks offer results from research on all of these instruments. The *Standards for Educational and Psychological Tests,* published in 1999, contains guidelines to judge the theoretical and psychometric qualities of these tests and questionnaires. Recent meta-analyses are a good source for judging the construct and criterion validity of many instruments.

In conclusion, all of the preceding instruments are based on the individual differences orientation. There are only a few instruments that use a developmental construct (e.g., Loevinger's sentence completion test to measure the level of ego development, Kerssies and colleagues' scales for sensorimotor development). However, these latter instruments are not frequently used by assessors. In the contextual orientation, the dependent variables usually are the well-known individual differences variables.

3.4. The Assessment Process in Children

The assessment process starts with the client's question (or problem) and ends with the advice (or solution). It is modeled according to the procedure of empirically testing hypotheses in a true experiment and contains the following:

- The problem/question as formulated by the client
- The assessor's reformulation of the problem in psychological constructs
- The assessor's formulation of several hypotheses containing different causes of the (problem) behaviors
- The choice of instruments to "test" the hypotheses
- Integration of all the information from dossiers, anamnesis, test results, and interviews with parents, teachers, and institutions into a coherent picture
- The professional advice or solution, including recommendations for a specific intervention, training, therapy, or program

Assessment differs from testing in that the latter is more oriented to measuring variables, whereas assessment is more oriented to answering a client's question or solving his or her problem. Testing is limited to the use of standardized instruments, whereas in assessment the judgmental processes and information integration play a significant role. Testing requires skills to conduct and score tests, whereas the assessor must also be familiar with the three components involved in assessing children. Assessment involves choosing, conducting, and interpreting tests and their results, the analysis of each specific case, and the weighing and integration of information. Finally, the assessor should be able and willing to account for all of the steps taken in this process.

Traditionally, pitfalls threaten the assessor. This issue concerns the integration of all the information from the various sources and is known as the clinical prediction versus statistical prediction controversy. In 1999, Garb reviewed all of these pitfalls and concluded that clinical prediction usually performs less well than statistical prediction. Pitfalls include the halo effect, fundamental attribution error, ignoring the base rate of phenomena, overestimating a salient detail and neglecting available statistical information, and giving information a wrong differential weighting.

It is recommended that the assessor use decision aids, such as the multi-attribute utility theory and the Bayesian rule, as much as possible to integrate information. A recent meta-analysis showed that clinical prediction is weaker than statistical prediction, but not to such a large extent as has been suggested. In addition, tests that traditionally demand much of the assessor's interpretive abilities, such as projective tests and interviews, are not as unreliable and poorly valid as is often

claimed. So, clinical prediction and controlled information gathering, using interviews and projective tests, and disciplined interpretation, using formulas whenever possible, can contribute to sound assessment of children's achievements and social adaptation.

4. POSSIBILITIES AND LIMITATIONS IN ASSESSING CHILDREN

Assessment is performed using the composition of three components in the assessment process. This section addresses the possibilities and limitations. First, theories and constructs offer the possibility of defining the core elements of children's achievements and adaptation but are often not specific enough for testing in a model. Second, the individual differences approach dominates the measuring of achievements and social adaptation, but more developmental theorizing that allows instrument construction is necessary to gain a more complete picture of children's achievements and socioemotional adaptation in the course of time. Even more constructs may be necessary to model the diverse and real developmental pathways. Third, the quality of test construction can be fostered using the *Standards for Educational and Psychological Tests,* especially the standards for reliability and validity of tests. Fourth, achievement test construction was usually controlled by practical questions and contained much practical wisdom, but it can be enriched by clear theoretical considerations about the structure of achievements and adaptations. Fifth, the assessment process was structured after the empirical cycle for conducting scientific research, but this process can be supported by models and decision aids to minimize errors. Sixth, the assessor should be sensitive to the pitfalls in choosing, integrating, and interpreting the diagnostic information.

The assessment process requires disciplined assessors, whereas clinical assessors sometimes act freely and creatively to do justice to their "unique" clients.

Educational and child assessment poses specific problems.

5. CONCLUSION

The assessor must be aware of the relevant three components and integrate these into his or her assessment process. The first component is biased toward measuring individual differences on one or more dimensions, and this may detract attention away from real development in the child's achievements and typical behaviors and from the influence of the context of these achievements and behaviors (e.g., events, stimuli, programs).

The assessor can be biased. Using formulas to integrate diagnostic information and standardized procedures in tests that usually allow for subjective interpretations helps to reduce such biases.

The assessor must operate in the complex environment of the child, parents, teachers, the institution, and the wider society, on the one hand, and in the psychological discipline, on the other. The latter offers much for accomplishing assessment but is so broad that the assessor must specialize in-depth. Recently, many meta-analyses have appeared and offer an equilibrated summary of the qualities of instruments and interventions. An assessment is best conducted by a team representing the knowledge and skills of several disciplines, thereby preventing individual biases.

6. CASE STUDY: AN EXAMPLE OF EDUCATIONAL AND CHILD ASSESSMENT

"Bryan" is a 10-year-old boy who is in the fifth grade. He has a younger sister who is in the fourth grade. Bryan's mother has no paid employment, and his father is a small businessman. The parents went to a school psychologist for advice because the teacher reported problems with Bryan's behavior in the classroom, and they wondered what was causing this. The teacher recommended that the parents seek professional advice for Bryan. The school psychologist listened to the parents and formulated their question as follows: "They want to know more about Bryan's lack of concentration and ask whether he himself experiences this lack of concentration. They would appreciate advice on how the teacher could help him." The psychologist reformulated the parents' question into the following diagnostic question: "How much are concentration and behavioral problems present, and how can these be explained?" This case was submitted to a professional institution, and it was decided to assess Bryan so as to give a professional and scientifically sound answer to the diagnostic question that

reflected the perceived problem of the parents, the teacher, and Bryan.

Bryan's problem was discussed with his parents and teacher. The parents had had no early problems with Bryan in starting to talk and walk. He cried a lot because he was eager to do more than he could. The preschool noted nothing special except that Bryan was inaccurate in fine motor tasks. He had to take the third grade of primary school twice because his reading and calculating were behind those of his classmates. Bryan could not concentrate well, and he quarreled a lot with the teacher. Some improvement was seen in the fourth and fifth grades, but concentration remained a problem and his writing was slovenly. In the eighth grade, Bryan said that he hated school and liked to stay at home. He often interrupted his tasks when sitting in front of the teacher, so he was placed at the rear of the class where he was quieter. The parents themselves said that Bryan needed order and structure, that he was willing to do his utmost, and that he was pleased with good results. He has broad interests. He has no steady friend, but this does not seem to bother him. He gets on very well with his father. Bryan is a scout, and the troop leaders consider him to be a "model child." The parents do not understand why he should need special help.

The teacher, on the other hand, called Bryan a chaotic boy who is not motivated to achieve. The teacher punished him, but this did not lead to any noticeable change in Bryan's behavior. His reading is average compared with his classmates, but it is actually 1 year behind the national average. His handwriting is bad, and he is forced to do many of his writing tasks twice. He makes too many errors in arithmetic because he is careless. He tries to finish his tasks as quickly as possible and then bothers other pupils. Sometimes he is very introverted. He is involved in conflicts in the schoolyard and when he works in a small group. He is not willing to admit that he makes mistakes. The teacher has suggested that Bryan be given professional help.

These complaints were clustered by the assessor into the following: lack of concentration, chaotic behavior, demanding attention, poor arithmetic, low reading level, bad reading and writing, low achievement motivation, conflicts with peers, and lack of social relationships at school. A first guess at the possible causes resulted in the following: inadequate visual organization, no planning, insufficient fine motor development, bad relationship with teacher, needs more structure, inefficient compulsion and punishment, low self-esteem, avoids contacts with peers at school.

Hypotheses as definite answers to the question were formulated, including criteria to test the answers:

1. Lack of perceptual organization, reduced concentration (WISC, observation, concentration test)
2. ADHD (DSM, 16-item ADHD questionnaire, observation)
3. Low self-esteem (observation, interview, two questionnaires)
4. Lack of achievement motivation (questionnaire)
5. Bad relationship with the teacher (interview, observation in the classroom)

The hypotheses were tested using a criterion set prior to the start of testing. The answers were as follows: 1, yes; 2, no; 3, no; 4, no; and 5, yes.

6.1. Summary of the Results for the Client (Bryan, Teacher, School, and Parents)

Bryan has an average IQ, his perceptual organization is more than 1.5 *SD* below average, his attention and concentration are more than 1 *SD* below average, his relationship with the teacher is very tense, and Bryan avoids going to school. His self-esteem and achievement motivation are average.

Advice for treatment was concrete, and its implementation was feasible. It was considered that it would be better if Bryan could change class for the next year. The relationship with the teacher had to be improved. A daily planning of tasks was made and agreed on by both Bryan and the teacher, and the tasks were divided into small units. Instructions were short and clear, a time limit was given for each task, and he could ask questions at fixed times. The teacher was informed about the assessment results.

6.2. Three Months Later

Bryan had started the new school year with renewed vigor. The relationship with his new teacher was normal. He still showed easily "off task" behavior, but this was simple to correct, and he ended his tasks within the time limits if supervision was available. His results were average, with only his progress in arithmetic still being slow; he had made progress compared with the past year. Bryan was more involved in peer activities and had a steady friend now. His parents said that Bryan was somewhat pleased about going to school.

6.3. The Psychological Investigation of Bryan and the Assessment Components

Concerning the instruments, all of the tests were constructed according to classic test theory, and results of IRT-based educational tests were used. The choice of these instruments was accounted for by the hypotheses; that is, definite testable answers to the questions were needed. The criteria for acceptance and rejection of the hypotheses were formulated in advance. There were several "rival" hypotheses to explain Bryan's problem behaviors; two of these were not rejected, and three were rejected. Although the rival hypotheses were not really independent, the process did contain steps for trying to discover the "real cause" of the problem. There was no ordering in the hypotheses; that is, there was an educated guess as to the most or least successful. The cause of the problem was sought mainly "in the pupil," although one hypothesis considered the relationship between Bryan and the teacher. There were more tests and questionnaires available to test the hypotheses. There was no argument for the specific choices (e.g., the quality according to *Standards for Educational and Psychological Tests*). Every instrument was about individual differences, and Bryan's results are compared with those of a norm group or a criterion. Projective tests, often used in socioemotional problems, were not used.

There was no concept of development involved in the assessment; however, there was a follow-up interview with the parents and Bryan's new teacher. There were nonspecific ideas about the effect of changing the classroom environment for Bryan. What was the exact condition changed (independent variable), and what was the precise behavior (dependent variable)? The process followed the prescriptions of the hypotheses testing model of assessment. Theory, as present in the instruments used, was scarce. Assessment was focused on answering concrete questions, but the theoretical base could have been stronger.

It is clear that this assessment protocol depended on the components of classic test theory and that IRT was used only because of the available educational tests. The individual difference orientation dominated, but implicitly a contextual orientation was present (changing the structure of the tasks with some effect on Bryan's behaviors), whereas there was very little developmental orientation. Theoretical possibilities for explanation of the behaviors were not structured and possibly were insufficiently exploited. The hypotheses were not ordered, and no real rival explanations of the problem behaviors were stated. The instruments were chosen because of the questions posed, and there was an attempt to use various methods to assess the same behaviors (e.g., tests, observation, interview). The assessment was discussed by a team, and the parents, the teacher, and Bryan himself were informed of the results.

The preceding analysis, using the comprehensive definition of assessment used in this article, helps to pinpoint the stronger and weaker elements in a concrete psychological investigation of a child. It is recommended that these components be used in conducting such investigations of clients.

See Also the Following Articles

Assessment and Evaluation, Overview ■ Behavioral Assessment in Schools ■ Child Development and Culture ■ Educational Achievement and Culture ■ Intelligence Assessment ■ Intelligence in Humans ■ Neuropsychological Assessment in Schools ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for ■ Psychophysiological Assessment

Further Reading

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological tests* (7th ed.). Washington, DC: Author.

Caspi, A., and Roberts, B. W. (1999). Personality continuity and change across the life course. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 300–327). New York: Guilford.

Cronbach, L. J. (1990). *Essentials of psychological testing* (4th ed.). New York: Harper & Row.

Fernández-Ballesteros, R., De Bruyn, E., Godoy, A., Hornke, L., ter Laak, J., Vizcarro, C., Westhoff, K., Westmayer, H., and Zaccagnini, J. (2001). Guidelines for the assessment process. *European Journal of Psychological Assessment*, 17, 178–191.

Garb, H. N. (1999). *Studying the clinician: Judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19–30.

Hambleton, R., & Swaminathan, H. (1985). *Item response theory: Principles and applications.* Boston: Kluwer–Nijhoff.

Jensen, A. J. (1998). *The G factor: The science of mental ability.* London: Praeger.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M.

(2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56,* 128–165.

Neisser, U. (Ed.). (1998). *The rising curve: Long term gains in IQ and related measures.* Washington, DC: American Psychological Association.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (pp. 703–733). New York: John Wiley.

Ter Laak, J., De Goede, M., & Brugman, G. (2001). Teacher judgments of pupils: Agreement and accuracy. *Social Behavior and Personality, 29,* 257–270.

Effective Classroom Instruction

Margo A. Mastropieri and Thomas E. Scruggs

George Mason University, Fairfax, Virginia, USA

Effective classroom instruction refers to the application of the "teacher effectiveness" variables, that is, those variables that have been demonstrated to bear the strongest relation to student achievement. These variables include time on task, content coverage, pacing, scope and sequence, questioning, feedback, and praise. Systematic application of these elements has been demonstrated to increase academic achievement.

GLOSSARY

allocated time on task The amount of time that teachers plan for instruction.

alterable variables The variables that teachers can actually alter (unlike age, gender, and personality) to improve student achievement.

engaged time on task The amount of time that students are actively engaged in instruction.

guided practice Student practice activities that are undertaken with direct teacher supervision.

independent practice Student practice activities that are undertaken independently.

pace The rate at which teachers proceed through instructional content; it also refers to the rate of presentation of information.

teacher effectiveness Teacher behaviors, identified through empirical research, that bear the strongest and most consistent relation with student achievement.

time on task The amount of time that students are working on instructional materials or attending actively to instructional presentations and discussions.

1. INTRODUCTION

Researchers over the years have been interested in what makes a "good" teacher. Early investigations focused on nonalterable behaviors such as personality, gender, intelligence, and even style of dress. Later, research focused on curriculum comparisons (e.g., phonics vs look–say methods of teaching reading) to determine effective teaching. Over the past several decades, researchers began to examine alterable variables, that is, those variables that teachers could alter in themselves to improve student achievement. This "teacher effectiveness" literature, although not without its critics, has been instrumental in identifying how teachers can best engage students in instruction to maximize their learning. Some important variables identified by teacher effectiveness research are provided in Table I.

Teacher effectiveness research has sometimes been criticized for its supposed emphasis on

TABLE I
Some Teacher Effectiveness Variables

Variable	Description
Time on task (engaged)	Direct student engagement with instructional presentations or materials
Planning for content coverage	Content presented to students during a given time period
Scope and sequence	Content to be covered and order of presentation
Objectives	Behavioral outcomes of instruction
Curriculum	Instructional materials and course of study
Pacing	Rate at which objectives are met
Teacher presentation	Teacher delivery of information, including elements such as structure, clarity, redundancy, and enthusiasm
Feedback and praise	Teacher acknowledgment and provision of information regarding performance and positive consequation of student responding
Guided and independent practice	Opportunities for student practice of acquired content or skills in supervised or unsupervised circumstances
Weekly and monthly review	Activities to promote long-term retention of academic content or skills

teacher-determined (rather than student-generated) curricula, transmission of knowledge by direct instruction (rather than construction of knowledge by student discovery and inquiry), and pointed emphasis on achievement test scores as outcomes (rather than measures of affect, classroom tests, performance or portfolio assessments, etc.). However, it has also been suggested that the "effective instruction" variables are important regardless of the type of teaching, learning, or assessment that is being undertaken. For example, even when curricula are student generated, learning is inquiry based, and outcomes are measured with classroom performance assessments, it is important that students be actively engaged in relevant activities and instructional materials. Likewise, regardless of content or approach to instruction, it is important that students receive relevant feedback and praise for their positive efforts.

2. OVERVIEW OF EFFECTIVE CLASSROOM INSTRUCTION

Effective teaching variables include those that involve planning for content coverage and that use effective teaching strategies. Planning for content coverage includes scope and sequence, objectives, curriculum, and pacing, whereas variables relevant to delivering instruction include maximization of academic engaged time, effective teacher presentations, appropriate use of practice activities, review, and formative evaluation. Making effective use of these variables can help to maximize the performance of general education students as well as that of students with disabilities or other special needs.

3. PLANNING FOR CONTENT COVERAGE

The importance of content coverage is obvious in that students almost certainly will not learn content or skills that have not been covered. Nevertheless, teachers vary greatly in planning for content coverage, and these differences are strongly related to student achievement. An important consideration in content coverage is that the amount and type of content covered must match the preskills and abilities of the students learning the content. Variables relevant to the general area of content coverage include objectives, scope and sequence, curriculum, and pacing.

Academic content is learned better if it is based on instructional objectives. Objectives state the desired outcomes of instruction so that the effectiveness of the instruction can be determined. Typically, an objective specifies (a) the content being addressed by the objective, (b) the conditions under which the student's performance is to be assessed, and (c) the criteria established for acceptable performance. For example, consider the following objective: "The student will identify 5 homonyms from a list of 10 word pairs with 100% accuracy." The content of the objective is homonyms, the conditions specify that students will be presented with a list of 10 word pairs, and the criterion for acceptable performance is identification (rather than production) with 100% accuracy. Another example of an objective is as follows: "The student will read four pages from the grade-level reading materials at a rate of 150 words per minute and will answer five comprehension questions with 100% accuracy." This

objective also specifies content, conditions, and criteria to be achieved.

Scope and sequence refer to the breadth and depth of content to be presented and the order in which the content will be presented. Scope and sequence considerations allow for long-term planning of instruction and can provide implications for time allocations and the overall pace of instruction through the curriculum. Scope and sequence of instruction can help to prevent instructional redundancy and can help to ensure that students have mastered prerequisite content or skills before learning new content. Most states have developed standards for learning across the curriculum areas as well as grade levels that are used to define the scope and sequences selected for use at the school and classroom levels. These state standards present the basic learning that is expected at each grade level in every subject area. For example, the state of Virginia has standards for learning from kindergarten through 12th grade in mathematics, science, English, history, and social sciences, and these subjects are then assessed on statewide achievement tests. These standards can be accessed on each state's education department Web site.

The curriculum includes the instructional materials used for learning as well as the course of study and the scope and sequence within each grade level. The curriculum serves as an interface between the student and the learning objectives. Curriculum materials do not determine what instructional objectives will be; however, when they are used appropriately, they support and enhance the meeting of instructional objectives. For example, a teacher stating "Today we'll cover Chapter 12" is not specifying an instructional objective but merely stating which curriculum will be employed. However, curricula do carry implications for how instructional content will be presented. For example, a "hands-on" science curriculum that includes materials for student construction of electromagnets and telegraphs implies a model of instruction that is very different from that of a curriculum that includes a textbook with a chapter on electromagnetism.

Curricula are most effective for a variety of learning needs when they are directly relevant to the learning objectives, are carefully sequenced, provide sufficient redundancy of key concepts, and contain relevant activities for teacher evaluation of student progress toward meeting objectives. The most effective curriculum materials are meaningful and relevant for all students, including male and female students and students from a variety of cultural backgrounds, and should

include adaptations (or suggestions for adaptations) for students with disabilities or other special needs.

4. USING EFFECTIVE TEACHING STRATEGIES

The most consistently identified teacher effectiveness variable is time on task. That is, the more time that students spend learning specific academic content, the better they will learn it. Distinction is made between allocated time on task (the time that teachers allocate to particular instructional activities) and engaged time on task (the time that students actually spend engaged in instruction). Although no consistent relation has been seen between allocated time on task and academic achievement, research has demonstrated that engaged time on task is positively related to academic achievement. Strategies for maximizing engaged time on task have been identified and include effectively managing transitions, avoiding digressions and other irrelevant verbalizations, and managing classroom behavior. Sometimes, teachers can improve student-engaged time on task by examining the amount of allocated time for class and monitoring their own teacher behaviors more closely. For example, teachers may find that they start classes 5 minutes late, end classes 5 minutes early, stop during the middle of class to search for relevant instructional materials, and stop to manage inappropriate classroom behaviors. Once such target areas are identified, teachers can design strategies to improve each area that will result in increasing the amount of student-engaged time on task during classes.

Effective teacher presentations have also been identified. Teachers maximize achievement when their presentations contain elements such as structure of lesson, clarity of teacher presentation, redundancy in emphasizing important concepts, and enthusiasm. Structure includes enlisting students' attention, providing a lesson overview that includes lesson objectives, providing outlines of the lesson and indicating when transitional points occur, and summarizing and reviewing key points as the lesson proceeds. It is important that students be made aware of the structure and objectives of the lesson so that they will know what is to be accomplished and how it will be accomplished. For example, teachers who begin with statements indicating the order of the learning activities (e.g., "First, we will …," "Second, we will be …," "Finally, we

will …") provide the structure for the lesson for all students. Clarity includes presenting clearly and directly to the point of the lesson, avoiding vague or unfamiliar terminology, and providing concrete understandable examples. All teacher dialogue should be directly pertinent to the objective of the lesson. Redundancy does not refer to unnecessary repetition; rather, it refers to reemphasis and restating of the key elements of a lesson, particularly significant concepts and rules. Redundancy refers to the provision of multiple opportunities to practice learning newly presented content. Additional practice opportunities can vary in format and length given that the goal is to provide students with many chances to practice learning new information. Many students not only benefit from redundancy but also require redundancy prior to mastering new content. Teacher enthusiasm helps to maintain students' attention and helps to model a positive attitude toward learning. Components of teacher enthusiasm include appropriate body movements and gestures, animated facial expressions, vocal inflections, positive acceptance of relevant student contributions, and a high overall display of positive energy.

Teachers also can increase achievement when they teach at an appropriate rate of presentation and when they maximize active student engagement with instruction or instructional materials. Selecting an appropriate rate of instruction not only is very important but also can be very challenging for teachers during the current era of statewide high-stakes testing. For example, if it is apparent that students have not learned a concept, it becomes necessary to back up and reteach that concept. This may require taking an additional class period or more depending on the students and the knowledge level required. The challenge for teachers is to ensure that all necessary content is covered before the end of the school year.

During teacher presentations, appropriate use of feedback and verbal praise is also associated with higher achievement. For example, praise that acknowledges student effort, specifies what the student did to merit praise, and demonstrates the association between effort and achievement is positively related to student achievement. Another important characteristic of praise is that it be genuine.

Following teacher presentation, guided and independent practice activities are associated with long-term learning and academic achievement. Guided practice activities are conducted under teacher supervision until it is certain that students are able to proceed independently. Students undertake independent practice activities when the students have demonstrated acquisition of skills or concepts being taught and need independent practice to further reinforce learning. With both guided and independent practice, instructional tasks and materials must be directly relevant to instructional objectives.

Overall, a model of effective instruction for an individual lesson includes daily review, statement of objective, teacher presentation, guided practice, independent practice, and evaluation of learning objectives. Weekly and monthly review activities are also related to long-term learning and retention of important concepts.

5. PROMOTING EFFECTIVE INCLUSIVE INSTRUCTION

Research has supported the use of teacher effectiveness variables with students with special learning needs as well as with normally achieving students. In inclusive classrooms (i.e., classrooms in which students with special needs are included with normally achieving students), use of teacher effectiveness variables has been seen to increase learning for all students. In inclusive classes, awareness of the range of student needs is even more important, and consideration of the effective classroom instruction variables can guide teachers' instructional decision making.

Many of the teacher effectiveness variables have a direct impact on the characteristics of students with special learning needs. For example, because time is an important mediator of achievement, maximized time on task can directly address students who are having difficulty in assimilating academic content. Structure, clarity, and redundancy in teacher presentations can be directly beneficial to students who have difficulty in sustaining attention as well as to those who do not grasp new concepts quickly. Enthusiasm is helpful for students who are less motivated to succeed in school, and appropriate rate (or pace) of presentation of information and maximized engagement can be critical variables for those who must maximize efficiency to make adequate progress. Frequent review and evaluation can help to determine whether instruction was successful and when further intervention is needed.

Teacher effectiveness variables have also proven to be useful as a model for addressing specific learning problems. For example, if a student is having particular difficulty in mastering specific academic content,

teachers can consider whether certain variables can be adjusted. That is, teachers can increase the amount of engaged time on task, or they can adjust the rate of presentation or overall pace of instruction for meeting instructional objectives. Likewise, they can adjust their teacher presentations, the rate of presentation, or student engagement. Practice and review activities can also be adjusted. By attending to these alterable variables, student learning in inclusive settings can be addressed.

See Also the Following Articles

Academic Failure, Prevention of ■ Mathematics, Teaching of ■ Reading, Teaching of ■ Teaching Effectiveness ■ Writing, Teaching of

Further Reading

Algozzine, B., Ysseldyke, J. E., & Elliott, I. E. (1997). *Strategies and tactics for effective instruction*. Longmont, CO: Sopris West.

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, *51*, 5–32.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook on research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.

Carnine, D. (1976). Effects of two teacher presentation rates on off-task behavior, answering correctly, and participation. *Journal of Applied Behavior Analysis*, *9*, 199–206.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Larrivee, B. (1985). *Effective teaching for successful mainstreaming*. New York: Longman.

Mastropieri, M. A., & Scruggs, T. E. (1997). What's special about special education? A cautious view toward full inclusion. *Educational Forum*, *61*, 206–211.

Mastropieri, M. A., & Scruggs, T. E. (2002). *Effective instruction for special education* (3rd ed.). Austin, TX: ProEd.

Mastropieri, M. A., & Scruggs, T. E. (2003). *The inclusive classroom: Strategies for effective instruction* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Pressley, M., & McCormick, C. B. (1997). *Advanced educational psychology for educators, researchers, and policymakers*. Reading, MA: Addison–Wesley.

Rosenshine, B. (1997). Advances in research in instruction. In J. W. Lloyd, E. Kameenui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp. 197–220). Mahwah, NJ: Lawrence Erlbaum.

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376–391). New York: Macmillan.

Scruggs, T. E., & Mastropieri, M. A. (1995). What makes special education special? An analysis of the PASS variables in inclusion settings. *Journal of Special Education*, *29*, 224–233.

Sindelar, P. T., Smith, M. A., Harriman, N. E., Hale, R. L., & Wilson, R. J. (1986). Teacher effectiveness in special education programs. *Journal of Special Education, 20,* 195–207.

Elder Abuse

Michael A. Rodríguez

University of California, Los Angeles, California, USA

Donna Benton

Los Angeles Caregiver Resource Center, University of Southern California, Los Angeles, California, USA

GLOSSARY

biopsychosocial model Model that describes how the interaction between the social, physical, and psychological characteristics of individuals results in a physical, mental, and emotional health outcome that influences their future behavior.

elder abuse (1) Intentional actions that cause harm or create a serious risk of harm (whether or not harm is intended) to a vulnerable elder by a caregiver or other person who has a relationship of trust with the elder. (2) Failure by a caregiver to satisfy the elder's basic needs or to protect the elder from harm.

feedback loop Scheme by which outcomes are continually processed back to the variables that determined the outcome. The new outcomes result in new interpretations of the variables, which results in new outcomes.

financial abuse Misuse of another person's funds or theft of money, property, or possessions.

neglect A type of physical abuse that can be either intentional or unintentional. In terms of elder abuse, neglect is the failure to provide the basic needs of the elder or to provide goods or services that are necessary to avoid or prevent physical harm, mental anguish, or mental illness.

physical abuse An act that may result in pain, injury, and/or impairment. Physical abuse includes bodily harm, neglect by others, and medical mismanagement.

psychological abuse A deliberate act inflicted on another to cause mental anguish.

sexual abuse Forcing another to take part in any unwanted sexual activity.

stakeholders Those who are invested in the lives of a particular individual. For the elderly, stakeholders can be health care personnel, social service agencies, friends, and families.

transgenerational violence Abuse that is "passed down" or continually generated throughout generations of a family.

Elder abuse is a social and public health problem that affects over half a million abused elderly victims in the United States as well as their caregivers, family, and community institutions. Abuse results in physical, emotional, and mental angst for victims, some of whom are cognitively and/or functionally impaired. There are also societal economic repercussions of elder abuse due to the increased need and demand for greater health care and social services to assist the victims. Although health care and law enforcement participation have helped assuage the situation, elder abuse detection, assessment, and treatment remain challenging. Adding to the difficulty of combating elder abuse is the booming

geriatric population, which is predicted to increase the strain on caregivers and societal institutions. This stress may perpetuate greater incidences of elder abuse in the future. Currently, multidisciplinary teams of psychologists, physicians, social workers, and public health advocates work together to educate, prevent, and intervene earlier in potential elder abuse cases. This article discusses the ways in which professionals are attempting to ameliorate elder abuse, in addition to providing a brief background of the causes and effects of this problem. The discussion includes methods of identifying victims and abusers, assessment of abuse, methods of intervention, and factors to account for in treatment.

1. INTRODUCTION

1.1. Background

Elder abuse has been a phenomenon recognized in medical and social practice since the 1970s, when the terms "granny battering" and "granny bashing" were mentioned in literature from the United Kingdom. In the same decade, work on aging populations established the presence of similar incidents in the United States. Terms such as elder mistreatment, elder abuse, and battered elders syndrome have variously attempted to describe abuse against elderly individuals.

1.2. Definition

The Institute of Medicine defines elder abuse (or elder mistreatment) as (1) intentional actions that cause harm or create a serious risk of harm (whether or not harm is intended) to a vulnerable elder by a caregiver or other person who stands in a trust relationship to the elder, or (2) failure by a caregiver to satisfy the elder's basic needs or to protect the elder from harm. While the terms elder mistreatment and elder abuse are frequently used to describe this behavior, the term elder abuse will be used in this article.

1.3. Types of Abuse

Elder abuse occurring either in the home or in an institution can be described as one or more of the following types: physical abuse, neglect, psychological abuse, financial abuse, and sexual abuse. This article focuses on the issues of elder abuse as it occurs in the home.

1.3.1. *Physical*

Physical abuse is an act that may result in pain, injury, and/or impairment. Physical abuse includes bodily harm, neglect by others, medication misuse, and medical mismanagement. Bodily harm or physical assault can take many forms, such as beating, shaking, tripping, punching, burning, pulling of hair, slapping, gripping, pushing, pinching, kicking, and the use of physical restraints.

1.3.1.1. Neglect and Self-Neglect Neglect, which is considered a type of physical abuse, can either be intentional or unintentional. Intentional neglect is the deliberate failure to provide the basic needs of the elder. It also includes the failure to provide goods or services that are necessary to avoid or prevent physical harm, mental anguish, or mental illness. Unintentional neglect occurs when the caregiver is not knowledgeable about the elder's needs or when the caregiver is restricted in the care they can provide due to his or her own infirmities.

Self-neglect is a failure to engage in activities that a culture deems necessary to maintain socially accepted standards of personal or household hygiene and to perform activities needed to maintain health status. Self-neglect has been categorized as a type of physical abuse by some authors. The highest national percentage of reported cases of abuse are those of self-neglect.

1.3.2. *Psychological/Emotional Abuse*

Psychological and/or emotional abuse can occur independently or can be related to physical abuse such as neglect or other forms of abuse. It is a deliberate act inflicted on the elder that is intended to cause mental anguish. Psychological abuse may include isolation, verbal assault (name calling), threats that induce fear (intimidation), humiliation, harassment, ignoring, infantilization, and emotional deprivation.

1.3.3. *Financial Abuse*

Financial abuse is the misuse of an older person's funds or theft of money, property, or possessions.

1.3.4. *Sexual Abuse*

Sexual abuse involves forcing the elder to take part in any unwanted sexual activity, such as touching that makes the elder feel uncomfortable or photographing the elder person while he or she is changing clothes or bathing.

1.4. Epidemiology of Abuse

The National Center on Elder Abuse (NCEA) reports that in 1996 an estimated 551,011 persons aged 60 and over experienced abuse, neglect, and/or self-neglect in a 1-year period. The report also adds that there were four times as many new, unreported cases of elder abuse. Another study noted that more than 105,000 elderly Americans were victims of non-fatal violent crimes in 2001.

An anticipated increasing geriatric population is predicted to expand the elder abuse problem. According to the 2000 U.S. census, the number of individuals over the age of 65 has increased from 31.2 million in 1990 to 35 million in 2000. Population specialists predict that with these continuing trends, by the mid-21st century there will be more elderly people than young people in the United States. The increasing number of senior citizens creates greater dependency on the immediate and extended family and demand for health and social services to accommodate to the increased longevity and chronic medical care of the elderly. Higher costs of health care, inadequate support for caregiving, and personality conflicts between the elder and the family or caregiver can strain relationships and may contribute to the possibility of abuse.

2. LAW

Federal policymakers first addressed the issue of elder abuse through the Older Americans Act of 1965, which was followed by the Vulnerable Elder Rights Protection Program in 1992. Since then, states have developed their own laws governing elder abuse management. State elder abuse statutes have primarily been based on laws developed to address child abuse and spousal/intimate partner abuse. The primary focus of elder abuse laws in most states has been mandatory reporting statutes and follow-up investigational procedures. Forty-two states have mandatory reporting laws, and eight states have voluntary reporting requirements. Psychologists are mandated reporters in 29 states and are encouraged to report elder abuse in five states. The investigating agency and the scope of authority of the agencies varies from state to state, but the Adult Protective Services (APS), the Long-Term Care Ombudsman, and the local law enforcement agency are the most commonly recognized organizations to whom a report can be made about suspected or confirmed elder abuse. Penalties for not reporting elder abuse, the time frame for emergency reports, and the maximum length of investigation also depend on the state. It is recommended that interested parties become familiar with the reporting requirements specific to each state (http://www.elderabusecenter.org).

3. THEORIES OF ABUSE: RISK FACTOR PROFILES

The literature suggests that all older adults may be at risk of abuse. Reviews of the literature have found conflicting profiles of both victims and possible abusers, indicating that it is not fitting to eliminate any particular profile; nor is it appropriate to say that a typical profile exists. The following sections discuss factors clinicians should be aware of when working with older adults.

3.1. Impairment of the Elder

Alcohol abuse by the elder is one of the most common risk factors for elder abuse. Although research shows variations in the impact of alcohol abuse on elder abuse, it has been found that the victim's risk for elder abuse could increase up to 10-fold. Other commonly established risk factors for elder abuse include low cognitive abilities and physical functional impairment. Those with depression, dementia and psychiatric illness are more likely to be abused. Elders who have difficulty eating are also at greater risk of abuse. These impairments may lead elders to live with family members. Elders are more likely to be abused if they live with family but do not have a living spouse. This places women, who are more likely to be widowed, at a greater risk for abuse. However, women may also be abused when they care for a spouse who may have been or continues to be abusive.

3.2. Perpetrator Characteristics

Other risk factors of elder abuse have also been proposed. Anetzberger suggested that elder abuse is primarily a function of the perpetrator's characteristics. This means that the caregiver's problems, pathologies, and perceptions may be the key to understanding who becomes an abuser. Per G. Anetzberger, life stresses, financial problems, mental disabilities, and lack of empathy for older people with disabilities may render some caregivers "ill-suited for caregiving and, given the potential dynamics associated with

caregiving, can make them prone to inflict abuse" (2000, p. 48). Research from the Three Model Projects on Elder Abuse found that in many cases the abuser was emotionally or financially dependent on the victim. Substance abuse by the caregiver was also related to risk for elder abuse.

3.3. The Interaction between Caregiver and Care Receiver

The caregiver–elder relationship also affects the potential for abuse. Risk may increase if the caregiver perceives the care recipient as difficult, combative, excessively dependent, or any combination of these. Homer and Gilleard found that socially disruptive behavior by the care receiver was related to elder abuse. However, others have found no such relationship. The duration, type, and intensity of care needed, cultural values, and individual family expectations also contribute to the possibility of abuse. Among ethnic minority caregivers, burden may be less likely to be an issue expressed by caregivers. Some studies have found risk factors associated with non-white or ethnic minority populations.

3.4. Multi-Factor Clinical Assessments

Given the conflicts that have emerged in research, the prudent clinician will continue to use all of these factors in his or her assessments. Many clinicians often overlook the importance of understanding the caregivers of older adults. Clinicians should recognize that caregiving is a risk factor for abuse, and paying close attention to the caregivers' attitude toward their roles may be useful in preventing abuse from occurring or being overlooked. Therefore, they need to observe the dynamics of the family relationship when meeting either the caregiver or the care receiver.

4. CLINICAL MANAGEMENT OF ELDER ABUSE

Elder patients can self-refer or be referred for a psychological review by their primary care physician or other health care provider. Whether the health care provider refers elders for suspected elder abuse or for changes in emotional and/or psychological and cognitive status, it is necessary that psychologists always screen elder patients for incidences of abuse.

The theoretical model that guides clinical management is described more fully in a 2003 Institute of Medicine Report. This model is adapted from Engel's biopsychosocial model. According to the elder abuse model, the social, physical, and psychological characteristics of the elder interact with those of the caregiver or trusted other. The interaction between the elder and the caregiver or trusted other person occurs in an environment governed by the socioeconomic conditions of the involved parties, the level of economic dependency, and the normative expectation of other stakeholders (health care personnel, social service agencies, friends, and relatives). Also important to the interaction between the caregiver and the elder is the sociocultural context in which they live, which encompasses the institutional or organizational locus (such as nursing home, private household), race or ethnic group, and social network of the elder and caregiver. The presence of social ties of the stakeholders to the elder and the caregiver may serve as a monitoring control on their interaction. Literature suggests that absence of this social network increases the vulnerability of the elder to the risk of abuse.

The physical, mental, and emotional health outcomes of the interaction between the elder and the caregiver impact their future behavior. These outcomes continually feedback to and remold the elder's and caregiver's social, physical, and psychological characteristics that will produce new outcomes. The feedback loop may sometimes result in the occurrence of abuse, either as a one-time incident or as a progression of violent episodes. This process-oriented model can help expand understanding of the etiology of specific types of elder abuse, help develop suitable interventions, and eventually lead to a reversal of the process.

4.1. General Evaluation and Analysis

4.1.1. Approach

The person who elicits the elder's history during a clinical visit should be patient and tolerant. Elders who arrive at a clinic for assessment may speak slowly. It is important to talk clearly and slowly to an elder patient, as he or she may have hearing impairments. Care should be taken not to infantilize the patient and not to subscribe to ageist attitudes and myths about the elderly such as forgetfulness, senility, dependency, ineptness, unproductivity, and unattractiveness. Cultural and ethnic differences between the physician and patient must be respected and considered. Tact, belief,

and discretion are paramount to developing a trusting relationship with the patient.

4.1.2. Observation

Anyone who assesses an elder for signs of abuse needs to pay close attention to patient–caregiver interactions. Increased discomfort, silence or monosyllabic responses, and fear exhibited by the patient when the caregiver is present may be indicators of elder abuse. In such situations, it is advisable to interview the patient when the caregiver is not present. Patients' responses and body language need to be critically observed. Fear, anger, infantile behavior, agitation, rocking (in the absence of other motor diseases), and sucking are some behavioral responses that may indicate possible abusive situations in the patient's life. Other indicators include, but are not limited to, confusion or disorientation while providing responses, withdrawal, denial of events and happenings, failing to talk openly, and providing implausible stories. Physical signs provide vital clues to possible elder abuse incidences. Cuts, lacerations, bruises, welts, dehydration, loss of weight, and burns should raise suspicion of abuse.

4.1.3. History Taking

4.1.3.1. Past History Theories about elder abuse indicate the role of transgenerational violence as a high risk factor for violence against a senior adult. Past history of domestic violence in the life of the victim (either as the perpetrator or the victim) needs to be elicited with a nonjudgmental and sensitive approach. A history of past relationships, family dynamics, the number of household members, education levels, and available societal resources such as visits from friends, other relatives, and neighbors provide clues to detect isolation and incidences of neglect. History of substance abuse (alcohol and or chemical dependency), employment status, housing, and financial status provide insights into possible cases of dependency and potential abuse. Sexual history should also be considered.

4.1.3.2. Family History Death of a spouse or partner and history of alcoholism among family members are areas that need to be explored. Serious psychiatric disorders in the family might provide a clue to the level of caretaking and dependency that the elder might be responsible for in his or her home.

4.1.4. Screening

Detection of elder abuse can be accomplished by routinely screening for abuse. The American Medical Association urges every clinical setting to use a routine protocol for the detection and assessment of elder mistreatment. There are several instruments available for screening the patient and/or the caregiver. Table I contains a list of screening methods and instruments along with information on their measurement properties. Once a case is identified as suspected positive, an intervention plan has to be developed and applied with the consent of the patient. The following sections describe other assessments that aid in identifying incidences of elder abuse.

4.2. Assessment of Cognitive Status and Personality of the Patient

Psychological assessment of an elder should include a comprehensive assessment of age-related conditions such as anxiety disorders, depression, mood disorders, sleep disorders, sexual dysfunction, substance abuse, and personality. Especially important are those tests that assess for risk factors of elder abuse, such as alcohol abuse and depression. The American Psychological Association recommends the following measures for assessing the psychological and cognitive status of the elder patient: The Beck Anxiety Scale, the Hamilton Depression Rating Scale, and other oral and written forms of depression-assessing instruments such as the Geriatric Depression Scale (GDS). Some authors also recommend the Beck Depression Inventory or the Yesavage Depression Scales. The CAGE instrument assesses alcohol usage. Additionally, the geriatric version of the Michigan Alcoholism Screening Test can be specifically used with older adults. The Mattis Dementia Rating Scale or Cummings Inventory for Alzheimer's Dementia helps evaluate dementia.

Health care providers need to look for overt and covert clinical signs of the patient's psychological and cognitive status during evaluations. Overt signs such as crying, silence, and irritability indicate a possible depressive state or social withdrawal; they could also be learned passive behavior. Prior to performing any assessments, the elder patient should be familiarized with the tests and procedures. Elders should have appropriate assistive devices such as eyeglasses and hearing devices. For those who are non-English speakers, the tester should be able to converse in the language spoken by the elder or seek the assistance of a professional interpreter.

TABLE I
Elder Mistreatment Measures[a]

Measure	Summary	Characteristics	Properties
The QUALCARE Scale (Philips *et al.,* 1990a, 1990b)	Assessment of six areas: physical, medical management, psychosocial, environmental, human rights, and financial	Fifty-three-item observational rating scale designed to quantify and qualify family caregiving	Extensive psychometrics reported: Interrater agreement range: 0.79–0.88. Chronbach's alpha: 0.81–0.95 on six subscales
Hwalek-Sengstock Elder Abuse Screening Test	Assessment of physical, financial, psychological, and neglectful situations	Fifteen-item assessment screen for detecting suspected elder abuse and neglect	Discriminant function analysis: nine items identified 94% of cases. Three conceptual domains: violation of personal rights, characteristics of vulnerability, and potentially abusive situations
Fulmer Restriction Scale (Fulmer and Gurland, 1996)	Assessment of physical, psychological, and financial restriction of older adults	Thirty-four-item scale designed to elicit information regarding unnecessary restriction of the older adult	Chronbach's alpha: 0.78 Interrater agreement: 0.93 Available in Spanish

[a]Adapted from Bonnie, R. J., & Wallace, R. B. (Eds.) (2003). Appendix A: Elder mistreatment measures and studies. In *Elder mistreatment, abuse, neglect, and exploitation in aging America.* Washington DC: The National Academies Press.

4.3. Risk Factor Assessment

An assessment of possible risk factors by a home care team or social worker can help identify needs and possible resources. Incidences of substance abuse, the availability of a caregiver and/or a support system from the patient's family or friends, financial dependency on the patient by other family members, and caregiver stress are factors that influence the plan of action. Any course of action that is recommended by the psychologist should be relayed to the elder's primary care physician. Available options should be discussed with the patient and the social worker if possible.

4.4. Treatment and Management

Once suspicion of elder abuse has been confirmed, the next step is to establish a plan of action to minimize the abuse or effects of abuse. A safety assessment is performed to ensure that preventive measures are adapted to the elder's needs. The elder's physical appearance, home environment, and personal surroundings should be considered when determining the immediacy of danger. A multidisciplinary team of professionals should be involved in the diagnosis, management, and prevention of suspected elder mistreatment.

If the elder is in imminent danger, immediate action should be taken to remove the elder from the situation. One possible method of intervention is to admit the elder to the hospital. In the hospital, the elder will receive medical treatment and not be endangered. Less serious cases of suspected elder abuse can be dealt with through continuous monitoring of the elder. A safety plan may also be developed to ensure that the elder knows what to do if placed in a compromising situation. This may include packing a bag of clothes and keeping copies of important papers, keys, money, and other necessary articles to take if the elder needed to quickly leave his or her home.

If the danger involves the caregiver's burden, an intervention should be designed to lessen the stress of the caregiver through services including home health aides, adult day care, and other types of respite programs. Community resources should also be consulted for possible safety, educational, caretaking, and social support services.

When determining a plan of action, it is important to consider the capacity of the elder and the caregiver, due to the need for cooperation of both parties and the

desire to accommodate the individual to whom services are directed. One such method of intervention is psychotherapy. However, no specific psychological intervention is preferred for elders. Instead, treatment options are individualized, led by the nature of the problem, therapeutic goals, preferences of the involved elder, and convenience for the involved parties. During psychotherapy, the psychologist must remember to be culturally sensitive and respectful of the individual's race, gender, sexual orientation, and social class when assessing issues of mistreatment and formulating interventions. The psychologist should also be attentive to the sensory deficits of the elder, particularly hearing and vision loss, which may hinder communication. Last, psychologist should consider whether psychological symptoms are caused or exacerbated by underlying medical problems. This will enhance the patient–psychologist relationship and enable more effective and productive therapy sessions.

Another method of intervention is education. Elders can learn how to protect themselves and gain empowerment through education. Education may be used to assist psychological interventions by providing additional information regarding the rationale, structure, and goals of psychotherapy. Psychoeducation can also help families caring for cognitively impaired elders. It can familiarize the family with the nature of cognitive loss, problem solving for practical problems, and providing emotional support to cognitively impaired elders who have experienced abuse. Additionally, education teaches other professionals about aging and how to investigate suspected elder abuse. These professional can advocate for the safety of elders while providing psychological services to this population.

Physicians can also prescribe medications to elders to cope with the involved abuse. These prescriptions may enable elders to function well within society or address the repercussions of the abuse. If an elder is abusing his caregiver, medication may be used to normalize the elder's behavior and prevent him or her from doing further harm.

Patient/family cooperation should also be considered when creating a plan of action. For the plan to be useful, it must be implemented by the assisting family. Thus, interventions that consider abused elders' families and their shared values are needed. If the family of the abused elder disagrees with the chosen method of intervention, it is suggested that psychologists be sensitive to the family's opposition and work with them to create modifications to an intervention.

The key to success in each of these intervention methods is the follow-up process. To determine the efficacy of the proposed intervention, the elder or caregiver must be contacted to ensure that the elder's needs are being met, that a safe environment is preserved, and that wellness of the abused individual is maintained. If interventions are ineffective, physicians must modify their plan of action by re-evaluating the abused elder's needs. One reason that physicians may need to re-evaluate action plans is a change in the elder's access to resources. Interventions must consider the elder's standard of living. This will guarantee the potential for intervention use by the abused elder.

An additional note of concern involves identifying potential challenges to suggested interventions. Challenges may include the resistance of elders to disclosing their personal experiences of neglect or abuse. This resistance may be associated with feelings of embarrassment or shame involving their mistreatment. Elders may also feel uncomfortable with the notion of receiving mental health services for their abuse. Their reservations require physicians to attend to their needs through reassurance of confidentiality during scheduled meetings, by participating in active listening with the elders, and by actively engaging with elders and their caregivers. This may require the enlistment of an individual who has a trusting relationship with the abused victim.

One specific way to counter sentiments of resistance is to validate both the positive and the negative feelings expressed by family members and victims. This validation will cause resistance to slowly diminish as elders begin to feel that their physician is trustworthy and wants to help them. This gradual progression toward an open relationship will foster an environment suitable for the disclosure of sensitive material.

Intervention efforts may be impaired if the elder is cognitively unaware. Should the abused elder suffer cognitive impairment, it is necessary for psychiatrists or geriatricians to ensure that the elder receive adequate care from a reliable caregiver. If the caregiver is a possible abuse perpetrator, or if he or she is unable to handle the elder's impairment, additional respite care or institutionalization should be considered.

5. ETHICAL ISSUES

When developing interventions in elder abuse cases, the clinician needs to be cognizant of the basic principles of autonomy, justice, beneficence, and nonmaleficence.

5.1. Autonomy

Autonomy is the right of self-determination. It is the right to choose one's actions or course in life as long as they do not interfere unduly with the lives and actions of others. A clinician must respect the options and choices made by the elder in regards to his or her living situation and caregiver. If the patient decides to continue living in the abusive situation, written information regarding emergency assistance, a follow-up plan, and a safety plan should be developed in consultation with the victim. However, it is imperative to assess the cognitive status and mental stability of the elder and caregiver.

5.2. Justice

The principle of justice allows everyone the right to that which he or she is due. Some elder patients may not receive adequate care because the caregiver and/or family is exploiting the finances of the patient. Although the caregiver may use the patient's funds to care for the elder, any unequal distribution of resources that jeopardizes the health and well-being of the elder and that which is done against the wishes of the elder is a violation of the rights of the patient.

5.3. Beneficence and Maleficence

Any intervention that is developed to address elder abuse should do no harm to the patient. The psychologist developing interventions should consult the patient, the patient's doctor, and the social worker before deciding on a course of action. Decisions that raise ethical dilemmas should be weighed for the merit of potential good in relation to the potential harm. This involves recognizing the limitations of the available resources for the patient, the caregiver, and the cognitive and general health of the patient and the caregiver.

6. FUTURE DIRECTIONS

Psychologists and other public health personnel should educate the public about elder abuse. Culturally relevant materials and messages tailored to suit the needs of minorities conveyed in English and other languages will help inform the general public about the issue. Information for caregivers on their role and the availability of resources in the community will improve early detection and management of elder abuse. Advocacy for the rights of elders, especially in the area of mental health, to support their social and emotional well-being is necessary. Collaborative research with community organizations and academic institutions involved in research will help identify new areas for development.

7. SUMMARY

Forms of elder abuse include physical abuse, neglect, psychological abuse, financial abuse, and sexual abuse. These types of abuse may occur independently or coexist with another form of abuse and may be difficult to detect. When assessing for signs of abuse, a clinician should pay attention to visible marks as well as emotional and mental expressions of fear, anger, confusion, and sadness. The caregiver's circumstance, competency, and relationship to the elder should also be considered in the assessment. One difficulty in recognizing and diagnosing potential elder abuse is that its symptoms and risk factors may resemble otherwise normative age-related illnesses such as dementia and depression. Also, no typical profile exists for either the caregiver or the victim. When assessing for abuse, a range of psychological tests should be performed and extensive past histories and current circumstances should be elicited to gain a full understanding of the elder's situation. After diagnosis, possible methods of intervention include psychotherapy, psychoeducation, and separation of the elder from the caregiver. In formulating an intervention, the needs and abilities of the victim and caregiver should be taken into consideration, as must the rights of the victim. Plans of actions should be routinely re-evaluated to judge the intervention's effectiveness and should be adjusted according to changes in the circumstances of the victim and caregiver. The clinician should approach the process of assessment to intervention with patience for disabilities that the elder might have and without discriminating against factors such as age, culture, gender, and social class.

Acknowledgments

See Also the Following Articles

Aging and Competency ■ Elder Caregiving ■ Personality and Emotion in Late Life

Further Reading

Abeles, N., & The APA Working Group on the Older Adult Brochure. (1997). *What practitioners should know about working with older adults*. Washington, DC: American Psychological Association. Retrieved on December 3, 2003 from http://www.apa.org/pi/aging/practitioners.pdf.

American Medical Association. (2002). *Diagnosis and treatment guidelines on elder abuse and neglect*. Chicago, IL: American Medical Association. Retrieved on December 9, 2003 from http://www.ama-assn.org/ama1/pub/upload/mm/386/elderabuse.pdf.

Anetzberger, G. J. (2000). Caregiving: Primary cause of elder abuse? *Generations*, *24*, 46–51.

Beauchamp, T., & Childress, J. (1994). *Principles of biomedical ethics* (4th ed.). Oxford, UK: Oxford University Press.

Benton, D., & Marshall, C. (1991). Elder abuse. *Geriatric Home Care*, *7*, 831–845.

Bonnie, R. J., & Wallace, R. B. (Eds.) (2003). *Elder mistreatment: Abuse, neglect, and exploitation in aging America*. Washington, DC: The National Academies Press.

Dyer, C. B., Pavlik, V. N., *et al*. (2000). The high prevalence of depression and dementia in elder abuse or neglect. *Journal of the American Geriatrics Society*, *48*, 205–208.

Engels, G. L. (1977). The need for a new medical model: A challenge for bio-medicine. *Science*, *196*, 129–136.

House, J. S., Umbersome, D., & Landis, K. R. (1988). Structures and processes of social support. *Annual Review of Sociology*, *14*, 293–318.

Lachs, M. S., Berkman, L., *et al*. (1994). A prospective community-based pilot study of risk factors for the investigation of elder mistreatment. *Journal of the American Geriatrics Society*, *42*, 169–173.

Lachs, M. S., Williams, C., *et al*. (1997). Risk factors for reported elder abuse and neglect: A nine-year observational cohort study. *The Gerontologist*, *37*, 469–474.

Lauder, W., Scott, A. P., & Whyte, A. (2001). Nurses' judgements of self-neglect: A factorial survey. *International Journal of Nursing Studies*, *38*, 601–608.

Lin, N., Ye, X., & Ensel, W. E. (1999). Social support and depressed mood: A structural analysis. *Journal of Health Behavior*, *40*, 344–359.

The National Center on Elder Abuse at The American Public Human Services Association. (1998). *The National Elder Abuse Incidence Study Final Report*. Washington, DC: National Center on Elder Abuse. Retrieved on November 19, 2003 from http://www.aoa.gov/eldfam/Elder_Rights/Elder_Abuse/ABuseReport_Full.pdf.

Paveza, G. J., Cohen, D., *et al*. (1992). Severe family violence and Alzheimer's disease: Prevalence and risk factors. *The Gerontologist*, *32*, 493–497.

Pay, D. S. (1993). *Ask the question: A resource manual on elder abuse for health care personnel*. Vancouver, Canada: BC Institute Against Family Violence. Retrieved on November 13, 2003 from http://www.bcifv.org/pubs/ask_question.html.

Quinn, M. J., & Tomita, S. K. (1997). *Elder abuse and neglect: Causes, diagnosis, and intervention strategies* (2nd ed.). New York: Springer Publishing Co.

Rennison, C. (2002). *Criminal victimization 2001: Changes 2000–2001 with trends 1993–2001*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice. Retrieved on January 30, 2004 from http://www.ojp.usdoj.gov/bjs/pub/pdf/cv01.pdf.

Sandefur, R., & Laumann, E. O. (1998). A paradigm for social capital. *Rationality and Society*, *10*, 481–501.

U.S. Census. (2002). *United States Census 2000*. Washington, DC: U. S. Census Bureau. Retrieved on November 15, 2003 from http://www.census.gov.

Elder Caregiving

Allison M. Burton, Jung Kwak, and William E. Haley

University of South Florida, Tampa, Florida, USA

GLOSSARY

activities of daily living (ADLs) Basic activities necessary for a person to perform on a daily basis, specifically bathing, dressing, toileting, making transfers, continence control, and eating.

culture Worldview, values, norms, and behavior guidelines shared by a group of individuals that influence and form the group's way of life.

ethnicity A group's shared personhood based on a distinctive social and cultural heritage passed on from generation to generation.

formal caregiver Any person who provides assistance or care to another for financial compensation such as a nurse, a home health aide, or a social worker.

informal caregiver Any person who provides assistance or care to another without financial compensation such as a family member, a friend, or a church volunteer.

instrumental activities of daily living (IADLs) Basic activities necessary for a person to perform to be able to live independently, specifically managing money, using the telephone, grocery shopping, personal shopping, using transportation, housekeeping, doing chores, and managing medications.

Caregiving for older adults is a prominent occurrence in the daily lives of many Americans. It is a topic that has received considerable attention from a variety of sources, including research, popular media, and society as a whole. This article offers an overview of the current body of knowledge in the field, including but not limited to specific topics such as the stressors associated with caregiving, the impact of caregiving, and interventions that may be beneficial.

1. INTRODUCTION

This article provides an overview of important issues faced by family caregivers, a topic that is of increasing interest in our society. Although illness and disability can strike at any time in the life cycle and lead to care by family members, this article focuses on caregivers for older adults. It reviews the importance and growth of family caregiving; stress process models used to guide research and clinical practice; common stressors faced by family caregivers; the impact of caregiving on psychological, social, and health functioning; individual differences in response to caregiving; and psychosocial interventions for caregivers.

2. OVERVIEW OF ELDER CAREGIVING

As the population ages and an increased number of individuals are faced with the responsibility of caring for aging loved ones, the topic of elder caregiving becomes increasingly prominent as an area of research, clinical practice, and public policy. Broadly defined, caregivers are individuals who provide assistance of varying levels to other people who are unable to perform these tasks for themselves. Often the term "caregiver" is associated with a group of individuals who are thought of as homogenous; however, the open nature of the definition allows a high degree of variability within this group. One important distinction is that caregivers can be either paid (formal) or unpaid (informal). Formal caregivers are those who receive financial compensation for the services they perform (e.g., nurses, home health aides), whereas informal caregivers are those who volunteer their time to assist loved ones (e.g., family members, friends). It is estimated that between 80 and 90% of individuals with physical impairments reside in the community and receive the vast majority of their care from informal sources.

The cost of informal caregiving in the United States, including things such as lost wages, keeping an individual out of a nursing home, and performing tasks that otherwise someone would have to be hired to do, was estimated to be $196 billion in 1997, the year of the most recent national reliable estimates. This dollar figure is more than twice the amount estimated for nursing home care during that same year. Informal caregiving, although substantially less of an economic cost to the family members compared with institutional care, is often coupled with a number of other physical, psychological, and social costs to the individual (discussed in greater detail later).

Individuals choose to take on the responsibility of caring for a loved one for a variety of reasons, including economic conditions, perceived obligation, a cultural ethic that caregiving is expected, and a positive desire to provide care as a meaningful and worthy activity. Formal care for a chronically ill individual is very expensive, whether provided by professionals who come into the home, in an assisted living facility, or in a nursing home. Medicare and Medicaid rarely cover such services fully; thus, many families simply cannot afford such care. However even in situations where financial considerations are not the primary concern, many families consider it their duty to provide the necessary care and keep their loved ones at home. It is also common for caregivers in these situations to report that they believe they can provide better care than can staff members who do not have personal relationships with their loved ones.

3. STRESSORS OF CAREGIVING: WHAT DO CAREGIVERS DO?

Depending on the nature of the illness, caregivers may face a variety of daily demands. These can include impairments in self-care and cognition as well as development of changes in personality and behavior. Other stressors can include physical care demands and issues specific to end-of-life caregiving.

Self-care impairments include higher level instrumental activities of daily living (IADLS) and more fundamental activities of daily living (ADLs). IADLs include tasks such as managing finances, housekeeping, using transportation, and taking medications independently. Loss of independence in IADLs occurs relatively early in the course of progressive dementias but can also occur due to less serious medical problems (e.g., arthritis) that may limit mobility. ADL impairment occurs during middle and later phases of dementia and includes loss of independence in fundamental skills such as bathing, dressing, feeding, and continence of bowel and bladder. Family caregivers who assist with self-care impairments may provide only minor intermittent assistance (e.g., assistance in balancing a checkbook). Assistance with personal care, such as bathing and managing incontinence, is more psychologically stressful because it may involve embarrassing personal care. With assistance with ADLs and IADLs, families may face issues such as role reversal.

Cognitive impairment produces its own unique stressors. In addition to producing decrements in ADL and IADL functioning, dementia can lead to problems such as becoming disorientation, repeating questions, and forgetting to do tasks. Dementia also leads to behavioral problems such as wandering, agitation, and delusional behavior. Table I lists 10 patient problems reported to be the most stressful by a large sample of caregivers for older adults with dementia. Of particular note is that behavioral problems were reported to be much more stressful than self-care problems—even incontinence.

Caregiving for persons with terminal illness creates additional stressors. Beyond the actual tasks of caregiving, other issues include stressors that are more difficult

TABLE I
Patient Problems Reported as Stressful by Caregivers

- Dangerous behavior
- Getting lost
- Embarrassing behavior
- Waking the caregiver
- Agitation and restlessness
- Anger
- Suspiciousness
- Depression
- Incontinence
- Repeated questions

Source. Adapted from Haley, W. E., Wadley, V. G., West, C. C., & Vetzel, L. L. (1994). How caregiving stressors change with severity of dementia. *Seminars in Speech and Language, 15,* 195–205.

to define. One of these is witnessing the suffering of the care recipients. Caregivers may also be responsible for management of complex medical regimens.

Interaction with the formal care system can be another important stressor. Caregivers report that it is very difficult to negotiate care, manage in-home assistance, and deal with nursing homes. Research has also shown that when caregivers receive formal assistance, it does not decrease their involvement in or commitment to informal caregiving; instead, it simply supplements the informal caregiving activities.

A national survey of family caregivers conducted in 1996 found that the overall mean number of hours of care per week was 17.9; however, this number ranged from 3.6 hours for individuals whose relatives were mildly impaired to 56.5 hours for those whose relatives had significant physical and/or cognitive disabilities. In addition, the stage of the illness is a major factor in affecting how much care a person provides. Research examining spousal caregivers of persons with end-stage lung cancer in hospice has shown that an average of 120 hours per week of care is provided by these individuals. The actual services performed by family caregivers can vary from occasional shopping and household assistance for a frail elder; to assistance with bills and organization for an individual with early-stage dementia; to assistance with dressing, bathing, and ambulation for an individual who has suffered a moderate stroke; to complete responsibility for toileting, eating, and medications for an individual with end-stage cancer.

To get a better understanding of the broader picture of caregiving in the United States, the National Caregiver Survey was conducted in 1996. This study included 1500 family caregivers of individuals with both dementia and non-dementia diagnoses and provided very valuable insight into the differences between the two. The participants in this study had been providing care for slightly longer than 5 years regardless of diagnosis; however, the dementia caregivers reported significantly more hours per week in the caregiving role. In addition, a significantly higher percentage of dementia caregivers were providing care for each of the ADLs surveyed than were non-dementia caregivers. These findings suggest that the dementia caregivers in this study experienced greater demands while in the caregiving role, which apparently carried over into other areas of their lives. The dementia caregiving group reported greater impacts in terms of family conflict, caregiver strain, time for leisure and other activities, and mental and physical health problems.

4. INDIVIDUAL DIFFERENCES IN REACTION TO CAREGIVING

A number of studies have shown that high levels of stress experienced by caregivers of elders, in particular those with Alzheimer's disease or other types of dementia, often lead to a number of negative psycho-social and health outcomes such as feelings of burden, social isolation, depression, and health problems. However, although most caregivers endure immense stress, they vary in their ability to cope successfully. Researchers have found that it is not necessarily the objective level of stress experienced by caregivers that affects their mental and physical health; rather, it is actually a combination of factors, including context of the caregiving experience, individual appraisal of the caregiving role and competence in that role, coping style, and social support.

First, the context variables, such as the caregiver's gender, age, and health, the quality of the caregiver's relationship to the care recipient, and the length of time providing care, are believed to exert influence throughout the caregiving experience. Socioeconomic status, such as income and education, can affect caregiving outcomes, including morbidity and mortality. It is well established that there is a negative relationship between socioeconomic status and chronic morbidity and mortality, especially among older persons. Socioeconomic characteristics can exert an indirect effect via stress, social relations, and/or health status.

It is common for one individual to take on the role of primary caregiver, with women being the most likely to provide care either for themselves or for their own spouses. It is estimated that 70% of all family caregiving responsibilities are taken on by women and that more than 50% of all women provide this type of care at some point in their life courses. This is significant because studies that examine the effect of gender on the caregiving experience have consistently found that female caregivers may be at greater risk for negative mental health outcomes. Female caregivers tend to report higher levels of depressive symptomatology and anxiety, and lower levels of life satisfaction, than do male caregivers. Such gender effects seem to exist among both spousal and adult child caregivers.

The quality and nature of the relationship between caregivers and care recipients can also influence caregiving outcomes. It is generally recognized that positive social ties have a beneficial influence on general well-being. In caregiving situations, greater affection or intimacy with care recipients can reduce perceived burden or strain and can lower psychological distress among caregivers. Negative social interactions while caregiving, such as family conflict and criticism of caregivers, are a risk factor for greater caregiver depression. Other problematic situations include caregivers who suffer abuse from care recipients and estranged family members who are left with "unfinished business." Spousal caregivers generally spend a substantially greater number of hours per week in the caregiving role, do so for a longer period of time, and report more difficulty in sleeping, less energy, and more fatigue than do nonspousal caregivers. Spousal caregivers are also generally more willing to provide care even when care recipients need extensive care, and they are much less likely to place care recipients in nursing homes, than are adult children. In addition, spousal caregivers consistently have been shown to be in poorer health than noncaregivers. This finding is also shown for adult child caregivers; however, the increased age of spousal caregivers places this group at an even higher risk for personal physical health conditions. These findings, taken together, suggest that spousal caregivers are very committed to their roles and responsibilities, even at the expense of their own health. The vast majority of nonspousal caregivers are adult children for elderly individuals; however, African Americans are more likely to be cared for by members of their extended family than are their European American counterparts. In addition, the societal changes in marital status and number of

children per family have introduced an increasing number of friends, primarily older, never married, childless women, who care for each other, and this population of caregivers is likely to increase during the coming years as the baby boom generation ages.

A second important variable in the caregiving experience is appraisal. Caregivers vary greatly in their subjective reactions to the same circumstances. Caregivers who appraise the problems they face as highly stressful are at increased risk for depression, whereas caregivers who are able to find meaning and benefit from their caregiving experience, and who have high levels of self-efficacy in caregiving, are more resilient to the negative effects of caregiving. Caregiver appraisals are often influenced by caregivers' personalities, with a number of studies showing that neuroticism leads to negative appraisals and higher depression. Anger tends to be one of the most common negative affects among caregivers, especially the ones caring for Alzheimer's patients and those who have poor physical health. One common cause of anger in caregivers that can be addressed therapeutically is inappropriate attributions that care recipients' problems are intentional rather than being the effect of Alzheimer's disease, stroke, or other brain impairments. In contrast, caregivers with personality characteristics such as high levels of optimism may be protected from some of the negative consequences of stress.

Third, research has demonstrated that coping style and social support are critical in explaining the tremendous variability in physical and mental health outcomes for caregivers. The role of resources such as social support and coping strategy is to buffer the effect of stressors on the outcome. In general, informal support networks lessen the negative consequences of caregiving such as emotional distress, health concerns, and economic strain. It is important to note that social support can be viewed not only in terms of quantity but also in terms of subjective perceptions of the quality of support. Coping is considered another important buffer and is usually separated into two categories: problem-focused coping and emotion-focused coping. Caregivers may use problem-focused coping by seeking information about their relatives' conditions, pursuing medical advice or rehabilitative care, and/or arranging for assistance with caregiving duties. In contrast, emotion-focused coping by caregivers may include adaptive efforts such as finding meaning in adversity and distracting themselves from caregiving concerns as well as maladaptive efforts such as engaging in excessive rumination, denial, and/or substance abuse. Problem-focused coping is generally

adaptive but may reach limits when little or nothing can be done to improve the objective aspects of the care recipients' functioning. At such a point, the ability to accept limitations may be most adaptive.

5. RACIAL AND ETHNIC DIVERSITY ISSUES

The sociocultural adaptation of the stress and coping model suggests that the role of ethnicity is greater than merely being a status variable reflecting the disadvantaged socioeconomic status of members of ethnic minority groups. This view asserts that one's ethnicity and culture exert an influence on nearly all domains and stages of the stress process model, including who provides care, how caregiving demands are expected to be handled and shared among family members, how caregiving experience is interpreted and perceived, how people cope and deal with stress during caregiving, and what stress outcomes are expected. As to why caregiving experiences differ by race and/or ethnicity, some studies suggest that racial/ethnic differences in values and beliefs about aging and caregiving may be responsible for the observed differences. Although it is dangerous to generalize about racial/ethnic differences due to the risk of stereotyping, the existing literature suggests that there may be important cultural differences in caregiving; however, the individual must clearly be viewed in his or her totality and not just as a member of a racial/ethnic group.

The literature indicates that African Americans often hold different role expectations and attitudes about providing care and coping mechanisms compared with those held by European Americans. African Americans tend to hold more positive attitudes toward older adults and perceive cognitive and/or physical decline in health among elders as part of a normal process of aging. In terms of caregiving experience, African Americans report heavier workloads and more unmet needs than do European Americans; however, African American caregivers also report that they experience lower levels of stress and burden as well as higher levels of satisfaction. Moreover, African American caregivers report increased levels of religiosity in response to caregiving. They are more likely to use their religious faith and activities (e.g., prayer) as a coping strategy than are European American caregivers.

Although studies on Hispanic caregivers are limited and there is great diversity among members of this ethnic group, traditional values of Hispanic families influence their caregiving experience. One of these traditional values is familism, which emphasizes the importance of the family connection and interaction as well as the family serving as the core of the social support system. Thus, the kinship network is considered to be a very important source of social support for Hispanic elders. Family loyalty and filial piety among Asian Americans, and importance of relational networks among Native Americans, are known to be other important cultural values of ethnic groups that help to create a strong support system among them.

6. BIOPSYCHOSOCIAL IMPACT OF CAREGIVING

The beginning of this article noted that although informal caregivers ease the financial burden of long-term care for society, they do so at the risk of experiencing a variety of psychological, physical, and social consequences of caregiving. Research at this point has clearly demonstrated diverse effects of caregiving, but questions remain as to what degree and in which areas such effects occur.

One of the most consistent and well-researched relationships regarding the effects of caring for chronically or terminally ill individuals is the increased risk of depression. The availability of validated quantitative tools for evaluating depression has led to comparability across studies and confidence in the strength of the results. The actual rates of depression vary; however, most studies find that caregivers who are involved in dementia care, or in other extensive caregiving, are two to three times more likely to be at risk for clinically significant depression than are their noncaregiver counterparts. In addition, many caregivers experience subsyndromal or subclinical levels of depression. Caregivers are also much more likely to use psychotropic medications for mental health problems, such as depression and anxiety, than are noncaregivers.

To better understand the relationship between caregiving and increased risk of mental health problems, it is necessary for researchers to identify risk factors. Documented risk factors for increased distress during caregiving include being female, experiencing other stressful life events (e.g., bereavement), lacking social support, experiencing family tension, being younger, being in poorer physical health, ruminative thinking, and having lower education. Research has

also identified a number of factors that appear to be protective against negative mental health outcomes, including high levels of self-esteem and mastery, high ratings of optimism, high ratings of life satisfaction, and low ratings of neuroticism. The findings regarding patient diagnosis have been inconsistent, with some reporting that caregivers for individuals with Alzheimer's disease experience higher rates of depression and others reporting that the rates are comparable between diagnoses such as cancer and Alzheimer's disease.

An important emphasis of research in this area seeks to determine how long these changes in mental health status last, with a fairly large number of studies finding that elevated levels of depression are seen in some individuals for as long as 2 to 5 years following the loss of loved ones, although there is also evidence that some caregivers experience improvements in mental health when their caregiving ends. Additional longitudinal data are needed in this area to better understand what places an individual at risk for this type of chronic bereavement reaction.

Research on the physical health effects of caregiving has varied greatly in its sophistication in assessing health outcomes. A common assessment of physical health status in caregiving studies is a single-item self-rated health question, and more of these caregivers rate their health as poorer than do noncaregiving controls or population norms. Many studies in this area have used this single item as a proxy for physical health in the participants, and although research has demonstrated that this method can be reliable, more sophisticated methods are needed for success in this research.

There have been few longitudinal studies of the impact of caregiving on physical health. One large study comparing caregivers with noncaregivers of similar ages indicated that caregivers showed increased physical health symptoms over a 2-year period, whereas noncaregivers showed no increase over time.

Other studies have used more sophisticated measures of physical health. These studies indicate that caregiving can lead to increased health risk behaviors (e.g., drinking alcohol, smoking cigarettes), lower levels of physical exercise, changes in sleep patterns, decreased immune system functioning, altered response to the influenza vaccination, slower wound healing, increased blood pressure, and altered lipid profiles. There has also been interest in whether or not these physical health changes lead to increased mortality in caregivers, and this question was examined in the Caregiver Health Effects Study. This prospective, population-based cohort study followed caregivers for an average of 4½ years and reported that caregivers who experience strain while in the caregiving role showed a 63% increase in mortality as compared with caregivers who did not experience strain and noncaregiving controls. Additional work is needed in this area to identify the mechanisms underlying these physical health changes and increases in mortality; however, this research has been very valuable in demonstrating that caregiving can represent a major and clinically significant strain.

Caregivers commonly report that their responsibilities infringe on their personal and social lives. Caregivers often are unable to take part in leisure activities, vacations, and/or valued activities such as church attendance. Spousal caregivers become devoted to their roles and the responsibilities they entail, often at the expense of their own needs, including social interaction. They report spending less time engaging in social activities (e.g., having lunch with friends, spending time with grandchildren) and in personal hobbies (e.g., gardening, sewing). A small group of caregivers, particularly adult daughters, are sometimes described as the "sandwich generation"; in some cases, they are responsible for caring for their own children as well as for their parents. Some researchers believe that this may become an increasing phenomenon given the current trend toward women holding off on having children until later in life. These caregivers very commonly report significant changes in their employment status (e.g., decreased hours, extended leave, leaving the workforce altogether) as a result of time constraints due to the needs of their aging parents. The specific effects vary based on relationships to care recipients; however, the overall pattern appears to be consistent in that caregivers sacrifice their social and personal roles to accommodate the needs of their loved ones.

The cyclic relationship among mental health, physical health, and social engagement is a concern for health care workers and clinicians. Increases in depression and physical health problems can lead to decreased social interaction; however, the reverse is also true in that role constraints that inhibit social activities can lead to negative mental health outcomes, specifically increased depression. An additional concern is in the case of bereavement following a period of caregiving. The literature has demonstrated that low social support and mental and physical health problems during caregiving can lead to complications during bereavement. Therefore, if individuals have experienced these changes as a result of the caregiving experience, they are at risk. It is important to identify caregivers who are at risk for

these negative psychological, physical, and social outcomes so that they can be targeted for assistance before the complications progress.

Finally, it should be noted that caregiving can lead not only to negative impacts but also to benefits. Recent studies show that caregivers often report benefits from caregiving such as a sense of satisfaction, a feeling that they are repaying care they received when they were young, and feelings of mastery and accomplishment. These positive feelings can occur simultaneously with negative effects.

7. INTERVENTIONS FOR CAREGIVERS

Each individual in a caregiving situation can experience his or her responsibilities and role as a caregiver differently due to that person's unique characteristics and circumstances. Moreover, caring for loved ones can be a meaningful and positive experience. However, caregiving can also lead to negative psychosocial and physical outcomes such as depression, lowered level of physical health, and lowered sense of well-being. Thus, many caregivers are in need of psychosocial and instrumental support. With an increased knowledge of how various stressors and other caregiving characteristics interact with each other and lead to different outcomes, various intervention methods have been developed and provided to caregivers in different caregiving settings. These interventions have also been evaluated for their clinical significance and effectiveness in both statistical and practical ways. A number of studies that summarize, evaluate, and examine the strengths and weaknesses of currently available caregiving intervention research literature have been published. These review studies can provide further detailed knowledge and information on the progress of caregiving intervention research. However, this article focuses on providing an overview of the caregiving intervention research, including primary entities being targeted by the interventions, types of interventions, and effectiveness of these interventions.

In general, interventions need to be tailored to specific needs of caregivers who have diverse and unique characteristics and caregiving experiences. For example, a caregiver who is dealing with severe behavioral problems due to dementia might want to learn about the nature and course of disease progression and how to handle behavioral problems effectively. On the other hand, a caregiver who has recently placed a loved one in a nursing home might seek an intervention that can help him or her to cope with feelings of guilt. Therefore, caregiving interventions are designed to reduce different types of stress outcomes so as to meet unique individual needs of caregivers and their care recipients.

There are a few major domains of stress outcomes for which various intervention methods aim to improve. These outcomes include how caregivers respond to stress in terms of mental and physical symptoms, improved general sense of well-being and quality of life, and delayed institutionalization of care recipients. Symptoms that caregivers often exhibit in reaction to caregiving stress include elevated levels of depression, anxiety, anger, and hostility as well as lowered health status (often measured by self-reported health and blood pressure). Also, quality of life of caregivers is often measured broadly by life satisfaction, sense of burden, mood, and social support. Finally, other outcomes targeted for intervention include care recipient outcomes such as functional status and delay of institutionalization of care recipients and use of formal and informal support as well as services available to caregivers.

Caregiver interventions are focused on either improving quality of life and coping skills of caregivers or reducing the objective stressors such as amount of care provided by caregivers. These interventions include psychoeducational, psychotherapeutic, supportive, and service-based interventions. The psychoeducational intervention is often offered with other types of interventions and aims to improve caregivers' knowledge about themselves, care recipients, and the environment. This type of intervention focuses on teaching caregivers to develop skills in dealing and coping with stressors through providing comprehensive information on characteristics of disease process of care recipients, resources and services that caregivers can use, and other specific coping skills. Psychotherapy is another form of intervention that is often used in conjunction with psychoeducation involving individual counseling to caregivers by trained professionals. This type of intervention is usually designed to teach coping skills and problem-solving techniques with the cognitive behavioral approach.

Supportive intervention is another type, and its main goal is to build support systems or networks among caregivers and to create an environment for them to discuss and share problems and feelings related to caregiving. Thus, supportive intervention is focused on allowing participants (i.e., caregivers) to exchange ideas and strategies for coping with stressors and

difficulties. Service-based intervention programs usually provide respite and/or adult day care services. Respite and adult day care programs are designed for caregivers to take some time off from caregiving responsibilities by providing assistance with ADLs and/or certain activities for care recipients. Other types of intervention methods include providing memory training, activities, and medications to care recipients as well as relaxation training to caregivers.

Most of these interventions are multifaceted in their design and method, and they usually are offered in conjunction with other methods. This makes it difficult to attribute any specific component of intervention to the success of intervention outcomes. Moreover, the caregiver intervention literature shows that most of these interventions show some level of effectiveness, yet it appears that there is no single method that is clinically significant and effective across caregivers. However, in general, a combined form of individual educational and psychotherapeutic interventions tends to result in better outcomes than do group interventions. Moreover, psychoeducational and psychotherapeutic interventions are found to be effective in reducing

levels of depressive symptomatology as well as anxiety and anger. In addition to psychoeducational and psychotherapeutic interventions, behavior management training, stress management, and use of certain medications are known to be effective in increasing sense of well-being and quality of life of caregivers as well as in delaying nursing home placement of care recipients with Alzheimer's disease.

The majority of caregivers are likely to benefit from gaining knowledge about the diseases, caregiving tasks, and available resources. Some caregivers can obtain additional help by receiving training in coping and problem-solving skills as well as in behavioral management of care recipients. In addition, some studies have found that providing a mixed form of interventions that are given to both caregivers (e.g., psychoeducational intervention) and care recipients (e.g., medication, therapeutic activities) can also be effective. Furthermore, a fairly new approach to caregiving interventions involves using technologies such as enhanced telemedicine, computers, and the Internet. Table II provides an overview of some of the types of interventions described previously, including definitions and examples.

TABLE II
Definitions and Types of Caregiving Intervention

Type	Definition	Example
Psychoeducational intervention	This aims at improving knowledge of caregivers about themselves, care recipients, and the environment. This type of intervention focuses on teaching caregivers to develop skills to deal and cope with stressors.	Problem solving, coping skills training
Psychotherapeutic intervention	This involves individual counseling to caregivers by trained professionals. It is mostly designed to teach coping skills and problem-solving techniques with the cognitive behavioral approach.	Cognitive behavioral therapy
Supportive intervention	The main goal is to build a support system or network among caregivers and to create an environment for them to discuss and share problems and feelings related to caregiving.	Support group
Service-based intervention	Goals include facilitating caregivers' use of formal services, improving competence of care recipients, and delaying institutionalization of care recipients. Services provided include respite care for caregivers and drug and activity therapy for care recipients.	Respite care, adult day care, medication therapy, and memory clinic (for care recipients with dementia)

Source. Adapted from Schulz, R., O'Brien, A., Czaja, S., Ory, M., Norris, R., Martire, L. M., Belle, S. H., Burgio, L., Gitlin, L., Coon, D., Burns, R., Gallagher-Thompson, D., & Stevens, A. (2002). Dementia caregiver intervention research: In search of clinical significance. *The Gerontologist, 42,* 589–602.

8. CONCLUDING COMMENTS

Family caregiving is a topic that demands greater attention from researchers, clinicians, and those concerned with public policy. The aging of the baby boom generation may stimulate greater attention to the possibility of changes in public policy that provide more support for caregiving families. For example, in other countries, family caregivers may receive pay in recognition of their efforts, adjustments to their retirement benefits, vouchers that can be used to purchase community services, and other forms of support that recognize the value of family caregivers in preventing expensive and unnecessary institutional care.

See Also the Following Articles

Coping ■ Stress

Further Reading

Aneshensel, C. S., Pearlin, L. I., Mullan, J. T., Zarit, S. H., & Whitlatch, C. J. (1995). *Profiles in caregiving: The unexpected career.* San Diego: Academic Press.

Aranda, M. P., & Knight, B. G. (1997). The influence of ethnicity and culture on the caregiver stress and coping process: A sociocultural review and analysis. *The Gerontologist*, 37, 342–354.

Coon, D., Gallagher-Thompson, D., & Thompson, L. W. (Eds.). (2002). *Innovative interventions to reduce dementia caregivers' distress: A clinical guide.* New York: Springer.

Haley, W. E., Allen, R., Reynolds, S., Chen, H., Burton, A., & Gallagher-Thompson, D. (2002). Family issues in end-of-life decision making and end-of-life care. *American Behavioral Scientist*, 46, 284–297.

Haley, W. E., Roth, D. L., Coleton, M. I., Ford, G. R., West, C. A. C., Collins, R. P., & Isobe, T. L. (1996). Appraisal, coping, and social support as mediators of well-being in Black and White family caregivers of patients with Alzheimer's disease. *Journal of Consulting and Clinical Psychology*, 64, 121–129.

Kiecolt-Glaser, J. K., Marucha, P. T., Malarkey, W. B., Mercado, A. M., & Glaser, R. (1995). Slowing of wound healing by psychological stress. *The Lancet*, 346, 1194–1196.

Mace, N. L., Rabins, P. V., & McHugh, P. R. (1999). *The 36-hour day: A family guide to caring for persons with Alzheimer's disease, related dementing illnesses, and memory loss in later life* (3rd ed.). Baltimore, MD: Johns Hopkins University Press.

Mittelman, M. S., Ferris, S. H., Shulman, E., Steinberg, G., & Levin, B. (1996). A family intervention to delay nursing home placement of patients with Alzheimer's disease: A randomized controlled trial. *Journal of the American Medical Association*, 276, 1725–1731.

Pearlin, L. I., Mullan, J. T., Semple, S. J., & Skaff, M. M. (1990). Caregiving and the stress process: An overview of concepts and their measures. *The Gerontologist*, 30, 583–594.

Roth, D. L., Haley, W. E., Owen, J. E., Clay, O. J., & Goode, K. T. (2001). Latent growth models of the longitudinal effects of dementia caregiving: A comparison of African-American and White family caregivers. *Psychology and Aging*, 16, 427–436.

Schulz, R., & Beach, S. R. (1999). Caregiving as a risk factor for mortality: The Caregiver Health Effects Study. *Journal of the American Medical Association*, 282, 2215–2219.

Schulz, R., O'Brien, A., Czaja, S., Ory, M., Norris, R., Martire, L. M., Belle, S. H., Burgio, L., Gitlin, L., Coon, D., Burns, R., Gallagher-Thompson, D., & Stevens, A. (2002). Dementia caregiver intervention research: In search of clinical significance. *The Gerontologist*, 42, 589–602.

Sorensen, S., Pinquart, M., & Duberstein, P. (2002). How effective are interventions with caregivers? An updated meta-analysis. *The Gerontologist*, 42, 356–372.

Stephens, M. A. P., & Franks, M. M. (1995). Spillover between daughters' roles as caregiver and wife: Interference or enhancement? *Journal of Gerontology: Psychological Sciences B*, 50, P9–P17.

Teri, L., Logsdon, R. G., Uomoto, J., & McCurry, S. M. (1997). Behavioral treatment of depression in dementia patients: A controlled clinical trial. *Journal of Gerontology: Psychological Sciences B*, 52, P159–P166.

Weitzner, M., Haley, W. E., & Chen, H. (2000). The family caregiver of the older cancer patient. *Hematology/Oncology Clinics of North America*, 14, 269–281.

Zarit, S. H., Reever, K. E., & Bach-Peterson, J. (1980). Relatives of the impaired elderly: Correlates of feelings of burden. *The Gerontologist*, 20, 649–655.

Zarit, S. H., Stephens, M. A. P., Townsend, A., & Greene, R. (1998). Stress reduction for family caregivers: Effects of adult day care use. *Journal of Gerontology: Social Sciences B*, 53, S267–S277.

Emotion

Belinda Campos, Dacher Keltner, and Molly Parker Tapias

University of California, Berkeley, California, USA

GLOSSARY

anger An emotion elicited by the appraisal that a demeaning offense or injustice has been committed against one's self or one's own; anger communicates the intent to redress the perceived wrong.

anxiety An emotion characterized by threat and vigilant cognition that can impair performance when conditions are perceived as providing a challenge that exceeds capacity; low levels of anxiety can enhance performance, but high levels of anxiety can impair ability.

emotional expression Displays of experienced emotion in the face, body, voice, gesture, and gaze; expressions of emotion convey important information to others about the sender's current state, likely behavior, and enduring dispositions and help to shape social interactions in important ways.

happiness The broadest and most recognizable positive emotion; the display of happiness is characterized by the Duchenne smile, which involves movements of muscles around the mouth and eyes.

interpersonal Emotion processes that occur at the level of ongoing social interaction where two or more people are interacting; in this more social context, emotions coordinate social behavior and smooth the progress of social relations.

intrapersonal Emotion processes that occur at the level of the individual, including the cognitions, physiology, perceptions, and behaviors associated with individual response to environmental opportunities or challenges.

social function The particular problem that an emotion was designed to solve and how it does so by structuring coordinated social interactions; recent theory argues that emotions solve problems related to reproduction and the raising of offspring, the problems of group governance and cooperation, and the negotiation of social hierarchies.

Emotions are distinct, biologically based psychological entities characterized by short-lived coherent patterns of perception, experience, physiology, and communicative action. At both the intrapersonal and interpersonal levels, emotions shape experience and outcomes in important domains of life, including health, work, and school.

1. INTRODUCTION

What role do emotions play in human social life? Western philosophy has traditionally regarded emotions with suspicion, as a disruptive force. The Stoics advocated distancing oneself from an emotional life,

whereas more recently Hobbes and Freud believed that following natural inclinations would lead to uncontrolled emotion and societal havoc. Still others, such as Aristotle, took the perspective that at least certain emotions (e.g., anger, jealousy) should be controlled or moderated.

Mounting evidence and theory within psychology has led to the opposite view of emotion, that is, that emotions are well-organized coherent systems that enable humans to respond quickly and effectively to environmental and social problems and opportunities. This article outlines this latter perspective, termed a social functional approach, and delineates how social functional approaches lead to an improved understanding of how emotions influence health, work, and school.

2. DEFINING EMOTION

Emotion is easily recognized and difficult to define. Currently, emotion researchers define an emotion as a short-lived, biologically based pattern of perception, experience, physiology, and communication that occurs in response to specific physical and social challenges and opportunities. This definition helps to differentiate emotions from related phenomena. Whereas emotions are elicited by fairly flexible interpretations of stimuli and have specific intentional objects, moods have less specific causes and endure for longer periods of time. Sentiments are consistent emotional responses to objects that can endure for lifetimes and be passed from one generation to the next (e.g., intergroup hostility). Emotion-related traits, such as hostility and shyness, reflect consistent ways of responding emotionally to broad classes of stimuli. Sensations (e.g., pain, an itch) are not explicitly related to the individual's strivings.

Most emotion researchers now agree that emotions serve adaptive functions but differ in the questions for which they seek answers. Psychologists who study the structure of emotion systems consider the origins of specific emotions, the behavior associated with emotion experience, and the personal and social consequences of behavior prompted by emotion. Psychologists who study questions of communication and cross-cultural universality focus on facial displays of emotion, taking care to document the distinct muscle movements that characterize some emotions. Their counterparts who are more interested in cultural variation use ethnographic or lexical methods that emphasize unique aspects of emotional experience

such as emotion words with no ready translations outside of the society of origin. Researchers who study appraisals of events that lead to emotional experience use narrative methods to study the patterns of stimulus interpretation that differentiate emotions. These various lines of inquiry have led to remarkable discoveries during the past 20 years or so regarding the role of emotion in development, psychopathology, memory, and brain function.

3. MEASURING EMOTION

Emotions have many facets and can be measured in many ways. Researchers interested in the expression of emotion can capture facial expression with the Facial Action Coding System (FACS) developed by Ekman and Friesen or can rely on ethologically significant gaze, posture, vocal, and/or gestural activities. These methods are fairly nonobtrusive and allow researchers to study emotion outside of the laboratory. Researchers interested in emotion-related appraisal or the understanding of emotion have turned to narrative methods and interview techniques and elaborate coding schemes to discern the meaning processes that give rise to and differentiate emotions from one another. Categorizing emotion words into prototypes of an experience is a method used to understand how individuals conceptualize everyday emotion experience. Other researchers have recorded activity in the autonomic nervous system to ascertain when emotions occur and how they differ from one another, for example, by assessing heart rate, respiration pattern, blood flow to the periphery, and skin conductance. Finally, a great deal of research continues to rely on self-reports of emotion experience through the administration of either formal Likert-type scales, such as the PANAS (which measures the disposition to feel positive or negative emotion), or through more direct inquiry of emotion experience.

4. SOCIAL FUNCTIONS OF EMOTION

Darwin and James laid the foundation for the study of emotion in significant and enduring ways. In *The Expression of the Emotions in Man and Animals,* Darwin contended that emotions must be serviceable. That is, emotions should be functional adaptations that prepare

the organism to act and communicate a readiness to engage in a particular behavioral action. This preparation should be expressed through nonverbal displays that should be universal across human cultures. Darwin also argued that particular expressions of emotion should be directly connected to the behavior they prepare the individual to enact. Disgust, for example, is a serviceable habit that prepares an individual to expel harmful contents before they are ingested and whose display includes protrusion of the tongue. This thinking laid the foundation for the study of facial expression and functional accounts of emotion that are so widespread today.

In his 1884 essay titled "What Is an Emotion?," James extended some of Darwin's claims in proposing that each emotion should be defined by a specific physiological response. This radical notion fell out of favor but is now supported by studies indicating that amusement, anger, disgust, embarrassment, fear, love, surprise, and sadness all are associated with distinct physiological responses in the autonomic nervous system. More controversially, James also claimed that emotional experience followed rather than preceded emotion-relevant behavior. For example, crying would precede feelings of sadness. During the mid-20th century, this counterintuitive idea generated a great deal of interest among psychologists who prioritized cognitive and information explanations over motivational and emotional explanations of human behavior.

What distinguishes contemporary social functional approaches to emotion? First, a social functions perspective builds on Darwinian theorizing about emotion to consider how the challenges and opportunities of the human evolutionary environment led to the development of specific emotions. Importantly, this perspective is complemented by a consideration of how human cultural practices elaborate on human biology to influence emotion experience and meaning. Second, a social functions approach considers the effects of specific emotion processes at the intrapersonal level of individual experience and at the interpersonal level of ongoing social interaction among individuals. Thus, there is consensus that the experience of emotion guides information processing in functional ways, triggering specific memories and patterns of perception and categorization. The autonomic nervous system activity of emotion is thought to support specific kinds of action (e.g., fight or flight). And the communication of emotion conveys specific feelings, intentions, and dispositions, evoking specific and adaptive responses in observers. Finally, a

social functions approach is concerned with the consequences of emotion, emphasizing how emotions facilitate the successful navigation of the social environment.

In contrast to previous approaches that viewed emotions as disruptive forces, a social functional perspective regards emotions as evolved answers to the problems and opportunities presented by human social life across our evolutionary past. In this view, emotions are specific efficient responses tailored to address the problems of physical survival and social living. Emphasis is placed on the utility of emotion in the contexts in which it evolved, whether it remains useful in current environments or has become inappropriate or maladaptive. Framing emotion in this way generates a series of important and interrelated questions about emotion:

1. Why do humans have emotions?
2. How do emotions influence survival and adjustment?
3. What are the interrelated parts of the emotion system?
4. What are the beneficial consequences of emotion?

A social functions approach naturally directs attention to the role of emotion in navigating the complications of human social life. Diverse theorists converge on three classes of problems and opportunities that emotions were designed to solve. Several emotions, such as disgust and fear, help humans to meet the problems of survival, including avoiding predators and toxins. Still other emotions, including filial love, desire, compassion, and rage, help humans to meet the problems of reproduction that center on finding mates, maintaining monogamous bonds, and raising offspring to the age of viability. Finally, still other social emotions, such as gratitude, anger, embarrassment, and awe, help humans to meet the problems and opportunities related to group governance. More specifically, these emotions are thought to promote and maintain cooperative alliances and help humans to fold into social hierarchies and mete out justice and punishment. The article draws on this perspective to outline an understanding of seven emotions that will prove to be particularly relevant to health, work, and school settings.

4.1. Anger

Anger is historically regarded as one of the most disruptive emotions. Functional accounts of anger have shed light on the beneficial aspects of this misunderstood emotion. Anger has been shown to be elicited by the perception that an injustice has

been intentionally committed against the self by another. This attribution of blame for an unfair outcome motivates an individual to (a) seek redressive action and (b) signal the intention to aggress if redressive action is not taken. To this end, anger is received as a threat by the target. According to emotion theorist Robert Frank, anger has two beneficial consequences. First, anger prepares an individual to aggress in response to serious transgressions, if need be, by directing blood to the face and arms. More important, signaling that one is genuinely likely to aggress is threatening and reduces the likelihood that a given individual will be cheated or taken advantage of by another. Signaling honest anger may be all the self-defense that is necessary to motivate a target to take redressive action or flee. Indeed, incidents of anger leading to actual aggression or violence are infrequent relative to the emotional experience.

4.2. Anxiety

Unlike anger, anxiety is not always defined as an emotion. It has not been documented to have a distinct nonverbal display or to be accurately decoded. Instead, anxiety has been related to fear, either as a subset of the broader fear category or as a more generalized mood-like state. The experience of anxiety appears to be a concomitant of threat-related cognitions when an individual perceives a challenge in the environment that may exceed resources. Small amounts of anxiety have been shown to enhance performance on a challenging task, but profound levels of anxiety have been shown to impair the ability to perform.

4.3. Embarrassment

The emotion of embarrassment is a clear example of individual experience diverging from larger social consequence. Embarrassment is marked by a brief display consisting of a downward gaze, a brief non-Duchenne smile, and the movement of the hand to the face. It is felt in response to unintended inappropriate behavior such as social gaffes, pratfalls, and uncontrolled bodily functions. Calling someone by the wrong name and tripping in public are typical triggers of embarrassment. Although embarrassment is an unpleasant individual experience, the audience present for an embarrassing incident is frequently amused or sympathetic toward the target. For this reason, embarrassment has been posited to serve an appeasement function. By feeling and displaying embarrassment, a person signals

to others that he or she realizes the inappropriateness of the behavior and is asking to be absolved of blame. At the level of the social interaction, the display of embarrassment restores social balance. Studies have shown that individuals who display embarrassment are better liked than their nonembarrassed peers and show greater concern for social norms. Adolescent boys who display embarrassment, for example, are less likely to engage in delinquent behavior than are their less embarrassment-prone peers.

4.4. Fear

The emotion of fear prepares an organism to flee from threats that pose an immediate danger or are perceived to be likely to overwhelm the system. Displayed through raised eyebrows, opened eyes, and an open mouth, the biological basis of fear has been well documented. In contrast to anger, which sends blood to the arms, fear sends blood to the legs, facilitating a flight response. In studies of fear response, young monkeys exposed to either a snake or a flower for the first time expressed a fear of snakes immediately, suggesting that they are prepared to feel a fear of snakes. Interpersonally, humans have been shown to be sensitive to the fear cues of others, valuing the warning that fear displays can provide. When a facial expression of fear is presented among a large group of smiling faces, the fear face is identified more quickly than is a happy face among a large group of fearful faces. The work of LeDoux has connected fear to the amygdala, an almond-sized structure of the brain shown to process fear-related information. Amygdala damage is associated with deficits in perceiving and attending to obvious threats.

4.5. Happiness

The emotion of happiness comprises the most general and recognized level of positive emotion. Happiness is universally recognized and signaled by the Duchenne smile, a combination that lifts the muscles of the mouth and crinkles the skin around the eyes. Unlike other types of smiling, Duchenne smiling reflects an individual's internal positive emotional state. The disposition to experience happiness has been associated with higher levels of life satisfaction and a cognitive style marked by optimism and diminished concern with social comparison. There is also some evidence that displays of happiness promote positive social relationships, perhaps by evoking positive social responses in others. A recent study showed that

women whose college yearbook pictures showed Duchenne smiling married earlier, had more positive relationships with partners and children, and reported higher levels of life satisfaction up to 30 years later.

4.6. Love

Love is traditionally regarded as an enduring attitude of affection toward another or a style of relating to romantic partners, and recent research on love has found evidence of emotion-like properties. Love is signaled through a series of affiliation cues that generate movement toward the love object and has been theorized to promote commitment between love partners. Consistent with this theorizing, the self-reported experience and display of love correlate with constructive conflict resolution and perceived trust during moments of relationship threat. Feelings of love are also associated with a greater willingness to engage in prorelationship behavior, even at a cost to self-interest.

4.7. Sadness

Felt in response to a perceived uncontrollable loss, sadness is marked by eyebrows pulled in toward the nose and a downturned mouth. Initially, sadness is marked by crying and other intense displays of grief, but prolonged sadness shuts down the body. This shutting down of the body is theorized to conserve the body's energy until a time when a person is better able to cope with the loss. Sadness is associated with cognitions of having little control over personal outcomes. For example, individuals who attribute positive life events to external causes and regard negative life events as having been brought on by internal characteristics of the self are prone to episodes of sadness.

The preceding seven emotion states do not cover the range of emotions currently studied by emotion researchers, nor do they encompass the states whose status as emotions generates the most agreement. Instead, these seven states comprise emotions whose relevance to the applied settings of health, work, and school has been most clearly documented.

5. EMOTIONS AND HEALTH

The literature on emotions and health is vast but can be summed up schematically with two fairly robust claims. First, positive emotions promote positive health and a longer and more satisfying life. Second, negative emotions impair health and place individuals at risk for poor health outcomes ranging from depression to heart disease. These empirical summaries are based on studies of specific emotional states (e.g., fear) but more typically emerge out of studies of emotion-based dispositions such as anxiousness, hostility, and happiness.

The emphasis on the beneficial aspects of positive emotions notwithstanding, positive emotion routes to better health remain understudied relative to negative emotions. Rather, the disposition to positive emotion and the cognitive concomitants of positive emotional experience have been related to better health. For example, dispositional optimism, which is related to happiness, is associated with longer life in the face of terminal illness. The positive illusions of individuals who overestimate the likelihood of surviving cancer or a stroke appear to act as protective resources, prolonging life and helping individuals to find positive meaning despite illness.

The relationship of positive illusions to health has been the subject of lively debate. In one study, optimism was measured in men recovering from coronary artery bypass surgery. A man's dispositional optimism was related to the release of liver enzymes that promoted recovery, and these men were quicker to recover the ability to sit, walk, and return to work than were their less optimistic counterparts. Other studies show that dispositional optimists are more likely to engage in health-protective behaviors such as taking vitamins, eating a well-balanced diet, and exercising. Still, researchers have contended that happiness and positive illusions should negatively affect health, perhaps by promoting denial or underestimation of the likelihood that illness will happen. In support of this view, unrealistically high levels of optimism have sometimes been related to a diminished likelihood of engaging in health-protective behaviors. Nonetheless, positive illusions remain one avenue through which positive emotional states have been shown to lead to improved health outcomes.

Turning to the negative emotional dispositions, it has become increasingly clear that anger, sadness, and (more recently) shame can have deleterious effects on health. Although previous research connecting health outcomes to the disposition to experience anger or sadness has oversimplified the relationship, the connection exists. Prolonged episodes of negative affective states, rather than a brief episode of negative emotion, are chiefly responsible for the relationship between negative emotion and health. Small doses of

negative emotion, such as anxiety and guilt, are associated with engaging in preventive health behaviors and being willing to seek help. However, more enduring chronic affective states, such as depressed mood and chronic anxiety, have been linked to impaired immune functioning, coronary heart disease, memory decrements, and other major health conditions.

Similarly, the propensity to experience frequent and intense episodes of anger is commonly related to poor health outcomes. The association of "Type A" behavior with heart disease is well known in scientific and popular circles. More recent research has shown that one component of the Type A profile, the propensity to experience hostility and anger, is the principal mediator of this association. Studies of Type A individuals who are less prone to hostility show that the goal-driven persistence that marks Type A may actually promote compliance of rehabilitative regimens, whereas the hostile component of Type A predicts future coronary episodes.

Prolonged episodes of sadness are also predictive of heart disease, but the emotional experience of sadness and its related cognitions have been most directly linked to depression. An attributional style that blames the self for negative life events and perceives little ability to control or create positive outcomes has been associated with the development of depression. More directly, the tendency to ruminate on negative events by persistently directing thoughts toward negative events and possible negative outcomes leads to a greater likelihood of suffering depression. In one study of coping in the aftermath of a strong earthquake, individuals who reported ruminating about their quake experience during the weeks afterward also reported more depressive symptoms and were more likely to seek health services. This coping style is more characteristic of women, and indeed, women make up a disproportionate number of depression sufferers.

Depression is associated with both increased negative affect (e.g., sadness, anxiety, anger) and decreased positive affect. Administration of selective serotonin reuptake inhibitors (SSRIs) tends to decrease experience of negative affect and increase experience of positive affect in clinically depressed populations. These drugs produce parallel effects in normal populations. Nondepressed individuals administered SSRIs for 4 weeks were rated as less hostile by clinicians, self-reported less negative affect, and showed more affiliative behavior in a dyadic social interaction than did individuals administered a placebo.

The experience of shame and self-blame may uniquely relate to negative immunological changes. Relative to a control condition, individuals induced to feel shame by writing about a past traumatic experience for which they blamed themselves showed increases in proinflammatory cytokines, negative immunological response, and the stress hormone cortisol. Self-reports of shame, but not of guilt or other negative affect, specifically correlated with these physiological changes. A study of HIV-positive gay men found that shame-related cognitions, such as poor self-worth and excessive concern with social rejection, are associated with lower CD4 count (a negative immunological marker) and higher AIDS mortality.

The literature connecting the disposition to feel negative emotion to poor health is extensive, but emerging advances in this area suggest that parts of this story have yet to unfold. Guided by social functional analyses emphasizing the beneficial aspects of emotion, researchers have begun to examine how negative emotion experience might affect health in positive ways. One proposed path through which negative emotions may positively affect health is by promoting "toughness" or resistance to the harmful effects of stress. For example, mice exposed to repeated stressors at an early age showed faster patterns of both physiological arousal and recovery, suggesting an increase in the effectiveness of the stress response over time.

In a related vein, the disclosure and free expression of negative emotions have also been shown to promote health. According to Pennebaker, one source of ill health is a repressive coping style that is characterized by an effortful inhibition of emotional thoughts, feelings, and behaviors. In support of this view, repressive coping has been associated with the release of cortisol. Focusing primarily on trauma sufferers who have not shared their experiences with others, Pennebaker found that sharing traumatic experiences by either talking or writing in journals leads to reduced physiological arousal and fewer visits for health care during the ensuing months. Similarly, other studies have shown that confiding or disclosing traumas, even secondhand traumas, is associated with better immune system functioning.

Consistent with a social functional account of emotion, negative emotion is not harmful per se and can be an adaptive response to a situation. Grief, for example, is not associated with the negative immunological changes that characterize depression. Grief may be an adaptive response to loss, whereas depression is maladaptive. Studies of HIV-positive individuals who have

lost partners or close friends to AIDS find that depressed mood is associated with negative immunological changes and more rapid HIV progression, whereas high levels of grief without depressed mood are not.

6. EMOTIONS AND WORK

A great deal of modern-day social interaction takes place in work environments. Consequently, a consideration of emotion in the workplace is particularly benefited by an interpersonal perspective. In the workplace, individuals must initiate relationships, strive for distinction and status, seek cooperation, seal commitments, and optimally, be personally satisfied with their work lives. These endeavors are facilitated by the communicative aspects of emotion, where people observe and respond to each other's emotions in ways that often help to resolve or change the relational problems of everyday life.

How do emotions influence work satisfaction and productivity? As with health, positive emotion is also associated with beneficial outcomes. The expression of positive emotion has been shown to positively relate to both work satisfaction and achievement. In one study of 272 employees, the expression of positive emotion was associated with more favorable supervisor evaluations, higher pay, and more support from supervisors and coworkers 18 months later. A related issue, and perhaps a conceptually dominant one, concerns the contextual factors that support the experience and display of positive emotion. Research suggests that much of the variance associated with job satisfaction is attributable to a perceived sense of efficacy in carrying out a task. Perceived efficacy is related to lower levels of chronic stress and higher levels of subjective well-being. Individuals who are given tasks that challenge, but do not exceed, their ability may achieve a state of "flow" in their work, where interest is high and the tasks are engaging. In contrast, perceiving the challenges of the work environment to be overwhelming can lead to unhappiness and high levels of job stress.

Serious negative occupational outcomes are also closely tied to interpersonal emotion. Job burnout is a response to interpersonal stressors on the job, where an overload of contact with people results in changes in attitudes and behaviors toward them. Field research has established that burnout involves three dimensions, with each dimension highlighting different emotions. Anger and anxiety are expected to be high in the dimension defined as emotional exhaustion, that is, feelings of being emotionally overextended and drained by one's contact with other people. Reduced experience of empathy and compassion is implicated in the depersonalization dimension, defined as an unfeeling and callous response toward people on the job. Decreased experience of pride appears to be central to the third dimension, where feeling low competence and low successful achievement in one's work leads to a reduced sense of personal accomplishment.

Another area of study has examined required emotional displays such as expressing positive emotion toward customers by workers in the retail industry or negative emotion toward debtors by workers at collection agencies. Termed "emotional labor," the use of positive emotional expression has been shown to elicit liking and a desire on the part of the customer to reciprocate in kind. On the other end of the spectrum, debt collectors are pressed to convey irritation to debtors and adjust their expressed emotions in response to debtor demeanor. This emotionally difficult work can clash with the collectors' genuine feelings toward debtors. This use of emotion may be task-effective, but emotional labor has been associated with decreases in felt positive emotion for expressing workers, and some organizations openly encourage employees to develop cognitive reappraisal skills to better cope with the emotional demands of the job.

How do emotions help individuals to navigate the social interactions of the workplace? How do they help them to initiate relationships, manage hierarchy, and procure the cooperation of coworkers? Anger and gratitude play important complementary roles in promoting cooperative alliances and creating stable long-term relationships. The initiation of cooperative relationships is assisted by liking and gratitude, which motivate the initiation of bonds and reward cooperation, respectively. Expressed anger can punish actions that threaten to undermine cooperation such as inappropriately taking credit and overusing resources. Threats that induce fear rather than anger have been found to be successful in changing the behavior of others. However, threats risk inducing anger rather than fear, and eliciting a strong angry response might pose a worthwhile risk only when interest and positive emotional interaction have proven to be ineffective.

The formation and maintenance of cooperative alliances take place within an existing hierarchical structure, and successful management of hierarchy is important to workplace success. Research suggests that the emotions of high-status individuals are of

primary relevance for two reasons. First, the outcomes of subordinates are strongly tied to individuals at higher ends of the hierarchy in most work environments. Second, the emotional experience of high-power individuals has been shown to differ from, and exert an effect on, the emotional state of subordinates. High-power individuals are more disinhibited in their emotional expression, expressing both negative and positive emotions and being less attune to how their emotional expressions are perceived by others. Status is conferred and reinforced by displays of anger rather than sadness, in part because the display of anger is associated with greater personal competence. Over time, the emotions of lower status people have been shown to converge to be more similar to the emotional experience of high-status people. This convergence is proposed to help coordinate the thoughts and behaviors of relationship partners, increase mutual understanding, and foster social cohesion. In the workplace, this emotional cohesion is likely to foster stable, mutually satisfying hierarchies.

In a workplace, decisions must be made and courses of action must be agreed on by relevant parties. How do emotions influence workplace decision making? One means is through emotion-related appraisals. Happy people view the world optimistically, perceiving situations to pose low risks and expecting positive outcomes. Fearful people perceive the world as threatening and are vigilant about risks and negative outcomes. These emotion-specific appraisal tendencies produce emotion-congruent effects on perception, reasoning, and decision making. People who feel happy readily notice and remember positive information, including other people's happy expressions. Bolstered by the accessibility and salience of positive information, happy people tend to view risks optimistically and make more risk-seeking choices. They label people using more positive categorizations (e.g., "good worker") and are more reliant on mental short-cuts in their decision making, and this can lead to less thorough information processing. In contrast, sad people tend to notice and remember negative and sad information, memories, and facial expressions. They tend to process information thoroughly when making judgments, presumably because doing so might help to explain or resolve their underlying negative mood. Appraising the world as threatening, fearful people assess risks pessimistically and make more risk-averse choices than do happy people. However, not all negative emotions lead to more careful decision making or judicious risk assessment. Anger is associated with

certainty and feeling control over a situation, and angry people behave more like happy people, assessing risks optimistically and making risk-seeking choices. Perceiving the world as under their own and others' control, angry people are more likely to blame individuals than to blame situational factors for negative events.

7. EMOTIONS IN SCHOOL

A large literature has addressed the factors that undermine academic achievement and social adjustment in school. The subset of the literature that relates to emotion has chiefly examined how anxiety leads individuals to underperform in, avoid, and mistrust particular academic domains or school in general. At a more interpersonal level, research on peer relations suggests important roles for anger, fear, sadness, and embarrassment in social adjustment.

Anxiety is considered to be the major cause of student underperformance on examinations. Students who consistently experience severe levels of anxiety are proposed to have a trait tendency known as test anxiety that involves increased anxiety and stress when specifically taking tests. Test anxiety has been associated with reduced motivation, suppressed immune functioning, and impaired test performance. Several mediators have been proposed, drawing on appraisal tendencies associated with anxiety and fear. Feeling afraid heightens a person's vigilance for threats in the environment. Studies find that test anxiety can impair performance by increasing a student's susceptibility to threatening external distractors. Other studies find that internally generated anxious thoughts disrupt performance by reducing the available cognitive resources in taking the test.

Anxiety is also considered to be a major cause of stereotype threat, a specific type of academic underperformance. Stereotype threat occurs when a negative stereotype about a group of which a person is a member becomes personally relevant, usually in the context of an experience the person is having. Stereotype threat is the resulting sense that the person can be judged or treated in terms of the stereotype or that the person might do something that would inadvertently confirm the stereotype. At each of level of academic skill, students sharing certain social identities get lower scores than do other students. Female students are outperformed by male students in math at each level of skill. Relative to White students, African American, Native American, and Hispanic American students

underperform on tests measuring overall academic ability. The anxiety-related processes believed to underlie stereotype threat have been proposed to be a primary cause of these groups' impaired performance. Self-reported anxiety has sometimes been found to mediate these groups' impaired performance, although in some cases only for those individuals who are sensitive to the group stereotype. One study found that African Americans did not report more anxiety than did Whites during an intellectual ability test; however, African Americans' blood pressure increased significantly from baseline relative to that of Whites.

How can the anxiety of stereotype threat be mitigated? Over a decade of research has found that the performance of students vulnerable to stereotype threat is improved when stereotype threat is removed from the test-taking situation, for example, when students are told that a particular test produces no group differences in performance. Other means of reducing stereotype threat include creating an academic setting that communicates identity safety, for example, when a teacher tells minority students that he or she will hold them to high academic standards and believes the students can meet them. The importance of combating stereotype threat is evident in the link with negative outcomes in academic domains that are negatively stereotyped for an individual's group identity. Students experiencing stereotype threat tend to avoid the domain where the stereotype applies, self-handicap in that domain, and/ or situationally disengage their self-esteem and ability appraisals from performance on tests in that domain. The chronic experience of stereotype threat risks instilling devaluation of success in that domain, a disengagement of identity from that domain, and a distancing from careers relevant to that domain.

Other emotions play critical roles in another dimension of school, that is, psychological adjustment and the negotiation of peer relationships. Anger is an important component of conduct disorder and externalizing problems in children. Children's self-reported tendency to experience and express anger correlates with their "acting out" behaviors in normal and clinically diagnosed populations. Children who engage in antisocial behavior are rated by parents, teachers, and peers as being more hostile and angry. Anger-prone children are found to expect hostility from their peers and respond to ambiguous situations with hostile behavior, which elicits rejection from other children. In contrast, children experiencing internalizing disorders, such as depression and anxiety, tend to feel fear and sadness. For example, measures of depression in fifth graders are associated with self-reports of sadness as well as anger in the case of boys.

The social–moral emotions, those that encourage moral behavior, socialization, and adaptive assimilation into society (e.g., embarrassment, shame, guilt), appear to motivate children to avoid social transgressions. Less guilt, for example, is positively associated with externalizing problems such as antisocial behavior and is negatively associated with moral development. One study of the psychological adjustment of adolescent boys tested for the roles of anger, fear, and embarrassment by coding facial expressions during a structured social interaction (an intelligence test). Boys categorized by teachers as having externalizing problems displayed increased expressions of anger, whereas boys categorized as having internalizing problems displayed increased fear. Interestingly, both groups of boys displayed less embarrassment during the task than did nondisordered boys, indicating that embarrassment signals normal social–moral development.

8. CONCLUSION

The three domains of health, work, and school are important contexts where life outcomes are shaped. Emotional experience and disposition shape individual experience and social interaction in ways that influence these life outcomes.

See Also the Following Articles

Agreeableness ■ Emotion and Culture ■ Intelligence, Emotional ■ Interpersonal Attraction ■ Interpersonal Perception

Further Reading

Darwin, C. (1998). *The expression of the emotions in man and animals* (P. Ekman, Ed.). Oxford, UK: Oxford University Press. (Originally published 1872)

Ekman, P., & Davidson, R. J. (Eds.). (1994). *The nature of emotion: Fundamental questions.* New York: Oxford University Press.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions.* New York: Norton.

Greenberg, M. A., Wortman, C. B., & Stone, A. A. (1996). Emotional expression and physical health: Revisiting traumatic memories or fostering self-regulation? *Journal of Personality and Social Psychology*, *71*, 588–602.

James, W. (1884). What is an emotion? *Mind*, *9*, 188–205.

Keltner, D., & Haidt, J. (2001). Social functions of emotions. In T. J. Mayne, & G. A. Bonanno (Eds.), *Emotions: Current issues and future directions* (pp. 192–213). New York: Guilford.

Mayne, T. J. (1999). Negative affect and health: The importance of being earnest. *Cognition & Emotion*, *13*, 601–635.

Salovey, P., Rothman, A. J., & Rodin, J. (1998). Health behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 633–683). New York: McGraw–Hill.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Emotional and Behavioral Problems, Students with

Thomas J. Huberty

Indiana University, Bloomington, Indiana, USA

GLOSSARY

comorbidity The co-occurrence of two or more mental disorders in the same person.

externalizing behaviors Behaviors that are undercontrolled (e.g., hyperactive, oppositional, delinquent, aggressive, disruptive).

internalizing behaviors Behaviors that are overcontrolled (e.g., withdrawal, anxiety, somatic complaints, obsessions).

symptoms Behaviors that are inappropriate or excessive and are of clinical significance.

syndrome A cluster of symptoms that occur together in a consistent and recognizable pattern.

This article presents characteristics of students with emotional and behavioral problems within the conceptualization of child and adolescent psychopathology. Implications for children in clinical and school settings are included.

1. INTRODUCTION

Students who have significant emotional and behavioral problems necessitating professional intervention often are described as being "emotionally disturbed," "behaviorally disordered," "emotionally and behaviorally disordered," or some similar term. Although these labels are in common use, they are broad and ambiguous and offer little help in understanding the nature, extent, prevalence, and implications of these problems in children and youth. This article describes current conceptualizations of children's emotional and behavioral problems and implications for schooling and effective interventions.

2. CONCEPTUALIZATIONS OF CHILD AND ADOLESCENT PSYCHOPATHOLOGY

The development of research and knowledge in child and adolescent psychopathology has burgeoned during the past two decades or so. The current state of the research suggests that children's emotional and behavioral problems can be conceptualized as internalizing or externalizing. Internalizing problems are characterized by anxiety, depressed mood, somatic complaints, withdrawal, and similar behaviors. These behaviors

often are referred to as "overcontrolled" because the children show inhibited behaviors and spend much energy in controlling behaviors. These problems also are similar to what are commonly termed "emotional problems." Externalizing problems refer to behaviors such as impulsiveness, aggression, inattention, hyperactivity, and distractibility. These behaviors are termed "undercontrolled" because the children lack the ability to control the behaviors. They are often referred to as "behavior problems." Internalizing and externalizing behaviors can be conceptualized as being on a continuum where the range of behaviors shown may be from very withdrawn (internalizing) to highly active and perhaps even aggressive (externalizing). However, this range does not represent a true continuum because children may show many behaviors simultaneously, although some may predominate. For example, 25 to 40% of children with conduct disorder (CD) also are depressed, although the conduct problems may be observed more readily. Research has shown that the correlation between internalizing and externalizing patterns ranges from approximately .25 to .40. Moreover, some behaviors are associated with both internalizing and externalizing problems, for example, social skills deficits and academic difficulties.

When children demonstrate internalizing or externalizing behaviors of such seriousness that personal, social, and/or academic functioning is affected, they may be referred to mental health professionals. Children with externalizing behaviors are more likely to be referred for mental health care because they are more disruptive at home and in the classroom. Children with hyperactivity, aggression, and other externalizing behaviors account for approximately 50% of all referrals to mental health professionals. Inattention without hyperactive behaviors is seen most readily in the classroom. Children with internalizing problems are less likely to be referred because they generally are not disruptive and the problems are less likely to be detected easily by parents or teachers. However, exceptions to this trend do occur, for example, severe phobia or depression that leads to suicidal ideation or behavior.

Another important conceptualization of behavior is trait and state, which complements the internalizing–externalizing perspective. "Trait" refers to behaviors that are characteristic of children and are seen across a variety of settings and circumstances. "State" refers to behaviors that are a manifestation of the impact of the environment. The state–trait concept is important for helping to understand why children's behaviors vary across settings. In schools, for example, it is common for children to exhibit off-task behaviors in one classroom but be more attentive in another classroom. Thus, the difference in behaviors might not be due to a trait; rather, it may reflect factors such as teacher management skills (state). Therefore, interventions may focus on the teacher and the setting rather than on the children. Children who have a high degree of a trait, however, may be more likely to show the behaviors in specific situations. Spielberger has conducted extensive research on anxiety and anger as having state and trait dimensions. Anxiety and anger tend to be constant traits, placing children at an increased likelihood of becoming anxious or angry in specific situations (state). Interventions often are directed toward reducing the magnitude of the trait as well as helping children to cope with anxiety and anger-producing situations.

2.1. The Taxonomy of Mental Disorders

A taxonomy is an orderly classification of objects, people, or other phenomena according to their presumed natural relations with each other. A taxonomy classifies according to mutually exclusive criteria; that is, the criteria can apply to only one specific member of a class, making the class unique. The concept is most often used in the physical sciences, such as botany and zoology, but it is also applied to human behavior. The taxonomy of mental disorders most commonly used in the United States is the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). There are more than 300 possible diagnoses of mental disorders in the DSM-IV, and determining which one is the most accurate description of the behaviors presented to the clinician can be challenging. The primary reason for the difficulty in formulating clear diagnoses is the problem of comorbidity; that is, many behaviors shown by children are seen in more than one disorder. When more than one disorder exists within the same individual (i.e., not mutually exclusive), they are considered to be comorbid disorders. Comorbidity is the rule rather than the exception in children and youth, occurring in more than half of all cases referred to mental health professionals.

2.2. Behaviors, Symptoms, Syndromes, and Disorders

Children's emotional and behavioral problems typically are first noticed by parents or teachers, who may refer these children to a physician, a psychologist, or other mental health professional. Symptoms are behaviors of

clinical significance that interfere with functioning. When behaviors occur together to form distinct patterns, they are termed "syndromes." When syndromes occur with such intensity, frequency, or duration that personal, social, or academic performance is affected negatively, the children may be given a diagnosis of a mental disorder by a psychologist or physician.

As an example, consider the case of "Eric," a 6-year-old boy in the first grade. As a preschooler, Eric exhibited some behaviors of concern such as high activity level, impulsiveness, and noncompliance with parents' requests in and out of the home. However, the parents were able to manage these behaviors successfully, and they did not view them as problematic. In school, Eric was required to participate in groups, take turns, and follow directions. Instead of these prosocial behaviors, a pattern of noncompliance, difficulties in participating and cooperating, blurting out answers, talking without permission, and interfering with other children developed. These behaviors were considered symptoms of a problem. When the symptoms occurred together in an identifiable pattern, they comprised a syndrome that was of clinical and educational significance because it interfered with the performance of Eric and others. He was evaluated by a psychologist, who considered the behaviors/syndrome in light of current diagnostic criteria and assigned a diagnosis of attention deficit/hyperactivity disorder (ADHD)–combined type. In this example, individual behaviors became symptoms that occurred together to become a pattern (syndrome). Because the syndrome was determined to be at a clinically significant level of impairment, a formal diagnosis of a mental disorder of ADHD was given.

3. MENTAL DISORDERS IN CHILDREN AND ADOLESCENTS

The majority of mental disorders listed in the DSM-IV can be seen in children and adults. The concepts of internalizing and externalizing patterns are applicable because most of the disorders can be described in those terms. There are some disorders, however, that are considered to be applicable specifically to children and adolescents, for example, separation anxiety disorder (SAD) and CD.

3.1. Internalizing Disorders

The most common internalizing disorders in children and adolescents are the same types that are seen in adults, although there may be some differences in symptoms. Internalizing disorders include conditions such as major depressive disorder, dysthymia, and bipolar disorder. Anxiety disorders occur in approximately 5 to 10% of children, and the prevalence of depression is approximately 3 to 5%. Table I summarizes characteristics of anxiety problems.

Recent research indicates that bipolar disorder may be more common in children and adolescents than was previously believed. Bipolar disorder was formerly known as manic–depressive disorder and shows a range of mood symptoms from high-energy mania to severe incapacitating depression. However, it is less common than anxiety or depression. Another internalizing anxiety disorder is posttraumatic stress disorder (PTSD), which was initially identified in soldiers who experienced extreme combat exposure, which was followed by "flashbacks" of the events, difficulty in sleeping, and generalized anxiety. In children, PTSD symptoms can be seen as a result of experiences such as abuse, divorce, and other traumatic events, although flashbacks are less common in children. The only anxiety disorder unique to childhood is SAD, which is an unwillingness to separate from parents at a level inappropriate for the age or developmental level of the

TABLE I
Characteristics of Anxiety in Children

Cognitive	Behavioral	Physiological
Difficulties with concentration	Shyness Withdrawal	Trembling or shaking
Memory problems	Frequent asking of questions	Increased heart rate
Overreaction and "catastrophizing" relatively minor events	Frequent need for reassurance	Excessive perspiration
Worry	Need for sameness	Shortness of breath
Irritability	Avoidant	Dizziness
Perfectionism	Rapid speech	Chest pain or discomfort
Thinking rigidity	Excessive talking	
Hypervigilant	Restlessness, "fidgety"	Flushing of the skin
Fear of losing control	Habit behaviors such as hair pulling and twirling	Nausea, vomiting, diarrhea
Fear of failure		
Difficulties with problem solving and academic performance	Impulsiveness	Muscle tension
		Sleep problems

children. It is common for children to show some anxiety about leaving their parents, even during the early days of school. When the behaviors exceed what is expected developmentally and the children cannot separate in a manner typical of most children, SAD may be an appropriate diagnosis. The presence of SAD may indicate parent–child or family issues; for example, a parent may have difficulty in allowing a child to have increased independence and may encourage and reinforce the child for refusing to separate. A similar pattern may exist in cases where a child exhibits "school refusal" (sometimes referred to as "school phobia"). If there is no rational basis for a child to not want to go to school (e.g., being the victim of bullying), it may suggest parent–child or family problems.

Depression occurs in less than 1% of preschool children and increases during the school years from approximately 0.3 to 2.5% in prepubertal children. By adolescence, however, the incidence of depression increases to approximately 4.3%, with lifetime prevalence rates of major depressive disorder in adolescents ranging from approximately 8.3 to 18.5%. Clearly, depression is a major concern with children and youth and is associated with suicidal ideation and attempts. Depression is a leading cause of death in adolescents after accidents. Some of the major characteristics of depression are presented in Table II.

3.2. Social Correlates of Internalizing Patterns

Children who have internalizing patterns, such as anxiety and depression, may tend to withdraw from social interactions as a way in which to cope with chronic feelings of distress. Withdrawal, social ineptness, and avoidance of interactions due to low self-esteem may be evident. These children may be perceived as "lazy" and unmotivated, despite the fact that they likely want to have social relations but are uncomfortable in initiating interactions due to perceived incompetence and lack of social skills. There is some evidence that anxious and withdrawn children are at higher risk for being the victims of bullies at school and may demonstrate reactive behaviors such as school refusal, sleeping problems, inconsistent performance, lack of assertiveness, poor peer relationships, low self-esteem, and low self-efficacy.

3.3. Academic Correlates of Internalizing Behaviors

Children who show anxiety and depression tend to have academic problems that can lead to more distress, causing a cycle of problems that persist over time. These children tend to avoid academically challenging tasks, withdraw from classroom activities, have poor

TABLE II
Characteristics of Depression in Children

Cognitive	Behavioral	Physiological
Arbitrary inferences	Withdrawal	Hyposomina or hypersomnia
Selective abstraction	Lack of assertiveness	Increased or lowered appetite
Overgeneralization	Impaired social relations	Weight loss or weight gain as a function of change in eating habits
Catastrophizing	Lack of energy	
"All-or-none" thinking	Lowered performance at school	Fatigue
Magnification and minimization		Cries easily
Personalization of too many or inappropriate events	Loss of interest in hobbies, activities, and friends	
Absolutist thinking	Low self-esteem	
Dichotomized thinking, leading to inefficiency in generating alternatives and problem solving	Seclusiveness	
Attributional styles: internal, stable, and global or external, unstable, and specific	Low participation	
	Irritability	
	Reduced concern about personal appearance	
	Negative statements about self and others	

achievement, lack persistence on difficult tasks, and show signs of "learned helplessness." They also tend to worry about appearing academically inferior to others, worry about incompetence, do not use effective learning strategies, and blame themselves for academic failure.

3.4. Externalizing Disorders

The most common externalizing disorders in children and adolescents are ADHD, CD, and oppositional defiant disorder (ODD). Collectively, these disorders often are referred to as disruptive behavior disorders (DBDs) because they have the common characteristic of showing disruption in home, school, and other settings.

ADHD is the most common DBD and accounts for a large percentage of referrals to mental health professionals. It is shown in approximately 3 to 5% of children and is three to six times more common in boys than in girls. It is a developmental disorder that appears to have some neurological substrates present at birth. With adequate assessment procedures, approximately 50% of cases can be identified by 4 years of age, although many children, especially those who tend to be inattentive, might not be identified until they enter school. There are three subtypes: predominantly inattentive type, predominantly hyperactive–impulsive type, and combined type. The inattentive type tends to be more common in girls than in boys and is more likely to be noticed at home and in situations where high degrees of attention are needed such as in studying, reading, and following directions. The hyperactive–impulsive type is characterized by displaying impulsive behaviors, engaging in off-task behaviors at school compared with peers, acting "fidgety," making noises, bothering others, and engaging in similar types of behavior. A list of DSM-IV criteria for ADHD subtypes is presented in Table III.

ODD is characterized by a pattern of oppositional behaviors such as noncompliance with requests or directives of parents, teachers, and/or other adults. Most of these children are resistant and defiant to authority and may engage in disruptive behavior but do not demonstrate significant antisocial behaviors. However, boys and girls with ODD are at high risk for developing more serious problems, primarily CD. CD is a well-established pattern of oppositional and defiant behaviors and is accompanied by antisocial behaviors such as stealing, fighting, truancy, and bullying. There are two forms of the disorder: (a) early-onset CD, which occurs before 7 years of age, and (b) late-onset CD, which is first seen during the preteen or early teen years. As might be expected, children with early-onset CD tend to have

TABLE III
Symptoms of ADHD

Inattention	Hyperactivity	Impulsivity
Fails to attend to details	"Fidgety"	Blurts out answers
Difficulty in sustaining attention	Leaves seat without permission	Difficulty in taking turns
Does not listen to directions	Runs about or climbs on objects inappropriately	Interrupts others
Avoids tasks requiring sustained effort	Difficulty in playing quietly	
Often loses things	On the "go," seems driven like a motor	
Distractible	Talks incessantly	
Forgets things and events		

more serious problems, are more likely to get into trouble with adults who attempt to exercise authority over them, or come into conflict with legal authorities. If the behaviors in early-onset CD are not corrected by late childhood or early adolescence, the prognosis for positive outcomes is guarded at best. Many children with early-onset CD are at high risk for developing antisocial personality disorder as adults, many of whom come into repeated conflict with society, including a high divorce rate, legal difficulties, alcohol or drug abuse, and employment problems. Children with early-onset CD are at a higher risk for school failure and for dropping out of school before graduation. Adolescents with late-onset CD are less likely to engage in significant antisocial behaviors. Instead, they are more prone to behavior problems such as poor academic performance, sexual risk taking, smoking, and noncompliance. Many children who are given a diagnosis of CD have met criteria for ODD at some time in their lives but might not have been given that diagnosis. Although most children with CD have met criteria for ODD, the reverse is less common because many children who present with ODD remain oppositional and do not develop the more significant antisocial behaviors associated with CD. ODD appears to set the pattern for CD by evoking coercive, harsh, and inconsistent parenting behaviors that lead to repeated conflict and management difficulties.

ADHD shows comorbidity with ODD and CD, but the relations are varied. Boys with ADHD are at higher risk for developing CD than are boys without ADHD; this

appears to be the result of high comorbidity of ADHD and ODD. In boys, ODD appears to be a precursor to CD, but ADHD is not a precursor to CD. Much less is known about these associations with girls; the associations may or may not be similar to those with boys. ADHD begins during the preschool to early school years and, if accompanied by early-onset CD, leads to more significant behavioral problems through the elementary school years.

3.5. Social Correlates of Disruptive Behavior Disorders

Children with DBDs are at higher risk for developing poor relationships with typical peers and adults, including parents and teachers. Their behaviors tend to be annoying and unacceptable to others, leading to punitive reactions by adults and social rejection and ostracism by peers. If the pattern continues, these children are likely to be seen as "outsiders" to their peer groups and will seek out other children with similar patterns, leading to reinforcement of the behaviors and even greater social alienation.

3.6. Academic Correlates of Disruptive Behavior Disorders

Children with externalizing behavior problems or disorders are more likely to do poorly in school, have academic skill deficits, exhibit frequent off-task behaviors, have poor peer relations in the classroom, and be disruptive to the educational process. They also may show signs of frustration, inferiority, anger, and aggression. Angry and aggressive children appear to have a greater number of fears, concerns, feelings of being victimized by others' actions and provocations, and beliefs that others have negative attitudes toward them. These feelings distract angry and aggressive children from focusing on academic tasks and acquiring necessary learning skills. Much as with the tendency toward social alienation, children with these behaviors may cause teachers and others to be less willing to help them, thereby creating more avenues for academic failure. Continued failure is likely to occur, creating a cycle of frustration and alienation. These children may, however, show "bravado" by exhibiting confidence in their academic skills while, at the same time, harboring resentment and anger at teachers whom they believe are not interested in helping them and at their peers who reject them. They may project the reasons for failure onto others and show boredom and lack of interest in school-related tasks.

4. INTERVENTIONS

4.1. Programmatic Interventions

Children who have emotional and behavioral problems might need help at school to assist them in achieving as much as possible. In the 2001 U.S. surgeon general's report, it was reported that 70% of all mental health services provided to children are delivered in the school setting. In general, there are two avenues for providing interventions in the school setting. The first avenue is through general education such as might be provided by school psychologists, counselors, and social workers. These interventions are based on clinicians' evaluations of children's needs and may take the form of individual or group counseling, consultation, or parent involvement.

The other avenue for intervention is determining whether children are eligible for special education services. Public schools in all states are required to have special education programs and services for children whose behaviors interfere with academic performance. These programs and their associated regulations must be consistent with the federal Individuals with Disabilities Education Act (IDEA), which provides assurance that all children who have disabilities are properly identified, assessed, and given appropriate instruction and support. For children with emotional and behavioral problems, IDEA uses the term "emotional disturbance" (ED), which is defined as follows:

(i) The term means a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child's educational performance:
   (a) An inability to learn that cannot be explained by intellectual, sensory, or health factors
   (b) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers
   (c) Inappropriate types of behavior or feelings under normal circumstances
   (d) A general pervasive mood of unhappiness or depression
   (e) A tendency to develop physical symptoms or fears associated with personal or school problems
(ii) The term includes schizophrenia. The term does not apply to children who are socially maladjusted unless it is determined that they have an emotional disturbance.

Although not all states are required to adopt this specific definition, many use the same or similar

terminology (e.g., "emotional disability," "behavior disorder"). The IDEA definition is problematic, however, because many of the terms are ambiguous and are not operationalized (e.g., "inability to learn," "satisfactory," "inappropriate," "normal circumstances"). The school psychologist is the professional most responsible for assessment of a child suspected of having ED, although the final determination of eligibility for special education services is done by a case conference committee that includes the parents, the school psychologist, a counselor, teachers, specialists, and others.

The child must have a "condition" that contributes to impairment in educational performance, which most often includes both academic and social development. It is important to note that having a formal DSM-IV diagnosis does not ensure that a child will receive special education services. It is not uncommon for parents to bring a diagnosis from a physician or psychologist to school, asking for special education services, only to be told that the child is not exhibiting problems that are interfering with educational performance. In those cases, the child might not be found eligible for services, despite having a formal psychiatric diagnosis. Conversely, a child could be found eligible for services for students with emotional disturbance and not have a formal psychiatric diagnosis.

A particularly enigmatic provision of the IDEA definition is the "social maladjustment" exclusion clause, which indicates that such children might not be found eligible for services. There is no definition of this term, and this can lead to difficulties for school-based professionals. The term, as originally considered, was to apply to children who exhibited delinquent behaviors that were controllable. Many professionals tend to equate "social maladjustment" with externalizing problems and equate "emotional disturbance" with internalizing conditions. This distinction would suggest that children with oppositional and conduct problems and ADHD would not be eligible for services unless there were accompanying internalizing problems that negatively affected performance. Conversely, because social maladjustment is not defined, one could argue that children with either internalizing or externalizing problems are socially maladjusted. Therefore, operationalizing the social maladjustment criterion for purposes of establishing eligibility is problematic. However, children with externalizing behaviors may qualify for special education services under IDEA or for general education accommodations under Section 504 of the Rehabilitation Act if needed to benefit from their education. They are not automatically excluded from receiving special education services.

If a student is found eligible for special education services for emotional or behavioral problems, an individualized education plan (IEP) is written in collaboration with the parents, school professionals, and the student if appropriate. In the IEP, goals and objectives are written to address the problems that are affecting educational performance. Services may include counseling, behavioral interventions, consultation with teachers, working simultaneously with the parents, and/or other appropriate interventions. The goal of special education services is not to provide comprehensive mental health services but rather to assist the student in performing at an appropriate level and in gaining educational benefit. The content of the IEP specifies what those services will be and how the outcomes will be measured. Counseling services may be provided if they are made a component of the IEP, or they may be included as an additional service under general education.

4.2. Individual Interventions

Individual interventions for students' emotional and behavioral problems are numerous and cannot be discussed extensively here. In general, research shows that cognitive–behavioral, behavioral, and multisystemic interventions are among the most effective. A brief summary of the major types of interventions and their goals are presented in the following subsections.

4.2.1. Internalizing Patterns

The majority of internalizing problems are seen in anxiety and depression, which have several common characteristics. Cognitive–behavioral therapy techniques focus on the faulty cognitions and attributions and include components such as attribution retraining, self-talk strategies, self-control methods, self-monitoring, and self-reinforcement. For older students with anxiety, progressive muscle relaxation coupled with systematic desensitization techniques have been shown to be effective in reducing the physiological symptoms of anxiety and in lessening phobic behaviors. Children with depression also tend to be socially withdrawn and to derive relatively little enjoyment from their environment. In these cases, activity increase strategies and reinforcement for increased social interactions may be useful adjuncts to cognitively oriented therapies.

4.2.2. Externalizing Patterns

Children with externalizing patterns, particularly those with aggression, hyperactivity, and impulsiveness, are

best addressed with behavioral intervention strategies that emphasize providing consequences for behavior, including rewards and negative consequences. Behavioral contracts, reward systems, token economy methods, reinforcement, and punishments often are effective methods for increasing prosocial behavior while decreasing undesirable behaviors. Aggression replacement training is a method that includes both cognitive and behavioral techniques to address aggression problems. Including parents and teachers in behavior management training, behavioral contracts, and reinforcement programs can be effective in helping students to change behavior over settings. If negative consequences are used, they should be used sparingly and coupled with positive reinforcement. Although physical punishment (e.g., spanking) is used by some parents, research indicates that it has only temporary effectiveness, at best, in stopping a behavior and does little toward teaching prosocial behavior. Other strategies include positive behavioral supports, cognitive–behavioral strategies, and multisystemic therapy.

Children and youth with either internalizing or externalizing problems often have difficulty with social interactions, self-control, and self-regulation of their emotions and behaviors. Therefore, teaching social skills, self-control strategies, and anger management techniques using cognitive–behavioral and behavioral techniques often is necessary to help children function more effectively with peers and adults in the school setting.

4.2.3. Medications

Prescribing medications to address emotional and behavioral problems in children and youth is controversial. The majority of psychotropic medications used with children were developed for adults and then adapted to children with dosage adjustments. In general, the research on short- and long-term effects of medications with children is sparse. The exception to this pattern is stimulant medications used to treat ADHD and associated symptoms. Historically, stimulants have been used to treat ADHD and been shown to be effective in approximately 70% of cases with few short- or long-term side effects. Some of the newer medications to treat ADHD are not stimulants and show promise of being effective as well as of reducing side effects such as sleeping problems. In the author's experience, parents tend to be reluctant to agree to use medication with their children, whereas school personnel are more likely to advocate its use, particularly for behaviors such as inattention, off-task, and hyperactivity. It is clear from research and clinical practice that medications, when used appropriately, can be effective in helping to address emotional and behavioral problems. Medication is most often prescribed to treat ADHD in children. The research literature demonstrates that medication combined with behavioral, cognitive, and self-management interventions is more effective than medication alone for the treatment of ADHD. Many children with ADHD have concomitant social and academic problems that are not affected directly by medication, thereby requiring behavioral interventions. Medication should be neither advocated nor completely ruled out for all cases; rather, it should be considered with regard to the needs of the individual child. School personnel are in a particularly good position to observe and monitor the effects of medication and can be valuable collaborators with the prescribing physician.

See Also the Following Articles

Attention Deficit/Hyperactivity Disorders (ADHD) ■ Attention Deficit Disorders: School-Based Interventions ■ Behavioral Assessment in Schools ■ Behavioral Observation in Schools ■ Diagnostic and Statistical Manual of Mental Disorders ■ Emotion ■ Learning Disabilities ■ Mental Measurement and Culture ■ School Violence Prevention ■ Traits

Further Reading

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002a). Is American student behavior getting worse? Teacher ratings over an 18-year period. *School Psychology Review*, *31*, 428–442.

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002b). Ten-year comparisons of problems and competencies for national samples of youth: Self, parent, and teacher reports. *Journal of Emotional and Behavioral Disorders*, *10*, 194–203.

Achenbach, T. M., & Edelbrock, C. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin*, *85*, 1275–1301.

Breen, M. J., & Fiedler, C. R. (2003). *Behavioral approach to assessment of youth with emotional/behavioral disorders* (2nd ed.). Austin, TX: Pro-Ed Publishing.

Cichetti, D. (1998). The development of depression in children and adolescents. *American Psychologist*, *53*, 221–241.

March, J. (Ed.). (1995). *Anxiety disorders in children and adolescents.* New York: Guilford.

Sameroff, A. J., Lewis, M., & Miller, S. M. (2002). *Handbook of developmental psychopathology* (2nd ed.). New York: Kluwer Academic/Plenum.

Stoff, D., Breiling, J., & Maser, J. (Eds.). (1997). *Handbook of antisocial behavior.* New York: John Wiley.

Wenar, C., & Kerig, P. (1999). *Developmental psychopathology* (4th ed.). New York: McGraw–Hill.

Emotion and Culture

Batja Mesquita
Wake Forest University, Winston–Salem, North Carolina, USA

Ashleigh Haire
University of North Carolina, Greensboro, North Carolina, USA

GLOSSARY

action readiness An inclination to achieve a particular change in one's relationship with the social environment; action readiness is an essential part of emotional experience.

antecedent event Any event that evokes an emotional response from an individual.

appraisal An individual's evaluation of the meaning an event has for him or her; an appraisal of the situation is an essential part of emotional experience.

cultural model A culture's construction of reality that organizes relationships and illustrates what a person is and should be; cultural models refer to understandings as well as the ways in which life is arranged.

dimensions of emotions Specific parts that make up the overall emotional experience of an individual such as the way in which the person evaluates an event and the way in which he or she is inclined to respond to it.

facial expressions Patterns of facial movements that are associated with emotion.

physiological response patterns Specific patterns of physiological events, such as muscle activity and heart rate, that are associated with the experience of a specific emotion.

prevention focus A psychological state in which an individual's primary goal is the pursuit of positive outcomes and rewards.

promotion focus A psychological state in which an individual's primary goal is the avoidance of negative outcomes and consequences.

vocal expressions Nonverbal aspects of vocal utterances that are associated with emotion.

Do people in different cultures experience similar emotions, or are the emotions of people in other cultures different from one's own? This is the primary question addressed in the research on culture and emotions. This question has important practical implications for a variety of issues. For instance, can one accurately recognize the emotional tone of parent–child relations in other cultures? Do Western and non-Western business partners understand each other's emotional expressions correctly when they are at the negotiating table? Is it possible to counsel people from other cultures? Can one interpret the aggressive threats or the populist emotional addresses of the leaders of antagonistic countries in other parts of the world? The question of the cross-cultural diversity of emotions is important in light of any social relations across cultures. To understand the answers to these questions, one must first understand what research shows about cross-cultural similarities and differences in emotional experience.

1. ARE THERE UNIVERSAL EMOTIONS?

Many people believe that, regardless of one's culture, there are certain things that we all experience as humans. Emotion is one such example. Quite a few studies in the domain of culture and emotion have aimed to make precisely this point—that some emotions are "basic" to all humans. Researchers who take this approach believe that, regardless of the culture in which one lives, certain emotions are experienced by everyone wherever they may live. In addition, these scientists suggest that each of these basic emotions has unique features. For example, many people believe that anger is a basic emotion and that the events that cause it, and the ways in which an individual responds to it, are fairly similar for everyone. An assumption underlying this viewpoint is that each of these basic emotions is characterized by a unique experience of the emotion itself and that this unique experience is not directly measurable by scientists. However, characteristic features of each of these emotions are measurable, and when they are observed, they identify the emotion being experienced by the individual. Researchers who seek to prove the existence of basic emotions have attempted to establish the presence or absence of these emotion characteristics (and hence the emotions themselves) in various cultures. Features of emotions include unique facial and vocal expressions, behaviors, and physiological arousal patterns. Thus, for these scientists, the question is a straightforward one: Does a particular emotion occur across cultures or not? The search to prove the existence of universal (basic) emotions has yielded some evidence for cross-cultural similarity in expressive behaviors and physiological responses.

1.1. Facial and Vocal Expression Research

Most research in support of universal emotions has focused on facial expressions. These studies have made use of a simple but elegant method. People from various cultures are typically asked to look at a picture of a face showing a given expression and to associate it with one emotion word from a list. The list consists of words for basic emotions such as happiness and anger. Across all cultures, people recognize a small number of facial expressions as representative of several emotions. This is true even for people and cultures with minimal exposure to the West.

Recognition studies on vocal expression (modeled after the facial recognition studies) also have found evidence for universality in the recognition of certain types of emotion expression. Individuals from various cultures were asked to identify the vocal expressions of emotions in either a nonsensical phrase or a phrase with meaning. People are clearly able to identify characteristic vocal expressions for a number of emotions (e.g., anger, happiness) even if they do not have information about the content of the utterance.

Cross-cultural studies on facial and vocal expressions show that humans have several universal modes of communicating their emotions. Therefore, certain elements of the emotion repertoire appear to be part of the human makeup. This evidence shows that emotions cannot be completely culturally defined.

However, the solid evidence for universality in facial expressions does not rule out the possibility of cultural differences or prove the existence of basic emotions. For instance, even though people can recognize facial and vocal expressions of other cultures, they are better at recognizing the expressions from their own culture. Thus, people judging faces and voices have an in-group advantage. Clearly, culture plays an important role in an individual's ability to recognize the expression of emotions.

The established recognition of facial and vocal expressions is not necessarily due to a direct correspondence between these expressions and distinct emotions. The results of facial recognition studies definitely suggest that people are able to know something about a person's psychological state from facial expressions (just as they can from anything else that a person does), but there is some evidence indicating that people would not spontaneously conclude that a particular basic emotion is occurring and that people draw this conclusion only when confronted with a scale that forces them to choose among various emotions.

1.2. Physiological Response Patterns Research

A few studies have also measured the physiological aspects of the autonomic nervous system associated with so-called basic emotions. In these studies, some individuals from various cultures were asked to relive an emotion from the past, and some were asked to move specific muscles in the face in motions that are assumed to mimic basic emotions. These studies showed some

evidence for similarity in physiological response patterns produced for several emotions across cultures.

However, there is reason to use caution in interpreting these data as conclusive evidence for basic emotions. First, in Western studies of emotions, it has not been resolved that emotions have characteristic physiological states. Scientists have pointed out, for example, that the physiological experience of fear from a loud sudden noise differs considerably from the physiological experience of fear of failing an examination. In addition, researchers have shown that there are no differences in the heart rates of people reliving fear, anger, and sadness expressions. Therefore, if a few studies show cross-cultural similarity in physiology of certain emotions, this does not prove that the same physiological responses always reflect the same emotions.

A second reason to use caution in interpreting these results is that participants from various cultures do not always experience those physiological patterns as emotions. In one study, people were asked to move their faces into the facial expressions typical of various emotions. The physiological responses of Minankabau men in Indonesia corresponded with those of American men who made the same facial expressions. However, the Minankabau reported experiencing no emotions, whereas the Americans reported experiencing emotions. Therefore, even if the pattern of physiology is similar across cultures, the emotional experience might not be.

In sum, scientists who believe that there are basic emotions have searched for unique patterns of features that are characteristic of particular emotions across all cultures. The research on facial and vocal expressions and on physiological response patterns shows that there are universal features of emotions. However, one cannot conclude that emotions themselves are universal because (a) some other features are culture specific and (b) the feeling associated with certain emotion features might be different across cultures.

2. WHAT COMPONENTS OF EMOTIONS ARE UNIVERSAL?

During the past two decades or so, some scientists searching for universalities have shifted away from the perspective that there are universal emotions. Instead, these researchers take the position that the building blocks, or "features," of emotions can be the same cross-culturally and that research should be done to establish the similarity of the features of emotions across cultures.

Two classes of features of the emotional experience are appraisal responses and action readiness responses. Appraisal is an individual's assessment of the meaning an event has for him or her. For example, a student who fails an exam will probably feel unpleasant (e.g., sad, angry) and be aware that the event has negative consequences. Action readiness is an individual's impulse to behave in some way toward objects in his or her environment in response to the emotion being experienced. The student might feel like "disappearing" or condemning the professor, even if these reactions are not feasible.

For scientists who investigate the features of emotions, then, people of all cultures are likely to evaluate emotional events in terms of how they rate on certain features (e.g., pleasantness, expectedness, controllability, agency). Where an event falls on such dimensions is also thought to correspond universally to similar types of experience. For example, the appraisal that a situation is positively relevant to oneself is an important feature of a positive emotional experience to everyone regardless of culture; likewise, a situation judged to be negatively relevant to oneself universally implies the experience of a negative emotion. Another example is the aspect of agency (responsibility). An individual who attributes responsibility for an event to another person might feel anger instead of the sadness he or she might have experienced if no such attribution of responsibility were made. Therefore, the same appraisals are thought to lead to the same emotions cross-culturally. Similarly, action readiness aspects contribute to emotional experience. The tendency to aggress toward another person is a central feature of the anger experience, and the tendency to quit doing anything at all is central to sadness.

Although emotion component theorists assume that the building blocks of emotions are largely similar across cultures, they do not assume that people appraise or desire to react to situations in the same manner across cultures. In this way, emotional experience may differ across cultures despite some universality in the building blocks. The extent to which emotions are actually different across cultures is a question that must be resolved by future research.

2.1. Universal Dimensions of Appraisal and Action Readiness

There is some evidence that the same dimensions of appraisal and action readiness help to distinguish among emotions across cultures and that they do so at similar levels. The ability to distinguish among emotions

is, to a large extent (~40%), due to appraisal dimensions. Among the dimensions that appear to be universally discriminative among emotions are pleasantness, expectedness, novelty, and agency (responsibility and control). Similar levels of the ability to discriminate among emotions are accounted for by action readiness dimensions. Current evidence suggests a core of action readiness dimensions that include moving away from, moving toward, and moving against the objects of emotions. So, there is some evidence supporting the idea of universal appraisal and action readiness dimensions that make up and help to distinguish various types of emotional experience across the world.

It should be noted that differentiation among emotions on the basis of appraisal and action readiness is only a partial explanation of what distinguishes emotions from one another, and this is true across cultures. There remains the possibility that some important dimensions have not been included, and there may be other important components of emotional experience (in addition to appraisal and action readiness).

2.2. Appraisal and Action Readiness Dimensions and Emotions

To what extent are certain emotions characterized by similar patterns of appraisals? For example, are experiences of emotions in the category of anger across cultures associated with the same appraisals? Studies have suggested that a number of emotions (e.g., joy, anger, sadness) share patterns of appraisals that are quite similar across cultures. For instance, joyful situations have been cross-culturally characterized as expected, very pleasant, requiring no action, and enhancing self-esteem. However, in these studies, cultural differences in appraisal patterns, in addition to these core similarities, have been established. Thus, happiness can be appraised as good for social relationships in some cultures (e.g., United States) and as bad for social relationships in others (e.g., Japan). Current research shows that the cultural differences in appraisals are smaller than the similarities. However, there are some reasons why this conclusion might not reflect the true picture. First, most studies have used relatively homogenous samples such as college students from various countries. The differences among students who are exposed to education with Western influences are likely to be smaller than differences among populations in cultures that are less influenced by Western culture. Second, because appraisal dimensions have been suggested by Western theorists, there

may be a Western bias in the proposed dimensions. There is a possibility that non-Western cultures have important appraisal dimensions that have been missed so far. For example, several studies that included East Asian cultures have suggested the importance of the dimension of interpersonal connection, a dimension not typically included in current appraisal research.

Studies addressing the cross-cultural similarity in action readiness have been relatively scarce and less systematic. There is some evidence that a small set of emotions are universally characterized by core action readiness modes. Culture plays an important role, however, in lending specific contextual features that help to clarify differences in action readiness modes. For example, across many cultures, shame has been associated with the urge to "disappear." In some other cultures, however, shame is associated with aggressive tendencies when it is caused by the behavior of someone very close to the affected individual.

In sum, research has shown that certain emotions have cross-culturally overlapping similarities in their core features. This finding supports some degree of the universality of emotions. At the very least, it suggests that similar patterns of appraisal or action readiness exist across cultures, and so there are similarities in the features of emotions. A limitation to this approach is its nearly exclusive reliance on retrospective self-report studies. So, one cannot be entirely sure to what extent the results reflect the ways in which people in various cultures think about their emotions or the ways in which emotions actually differ. On the other hand, it may be argued that emotional experience does not happen independently from the way in which people think about it and that self-reports provide a window on this experience that cannot be obtained in any other way. An additional limitation to this approach is that researchers of emotion features have been more interested in predicting similarities than in predicting differences. The differences that exist are equally informative about the nature of emotions and have been addressed by another group of researchers.

3. ARE EMOTIONS CULTURE SPECIFIC?

During the past two decades or so, some scientists have approached the issue of culture and emotion quite oppositely. Instead of looking for similarities in emotions and their components, these scientists have focused on the

cultural differences in emotions. These culture researchers believe that emotional experiences can be truly understood only by taking into consideration the culture in which they occur. Thus, some researchers try to understand how individuals' emotional responses relate to the values and practices of the culture.

The search for differences in emotions among various cultures begins with an insight into the ways in which cultures organize relationships and understand what a person is and should be. These social arrangements and understandings are referred to as cultural models. For example, achievement is one prevalent middle-class American cultural model. The belief that one should stand out from the crowd, the value placed on positive rewards, and the emphasis on personal influence and responsibility all are integral parts of this cultural model. Culture researchers have established that the particular models that are most important differ according to the culture in which one lives. For instance, this American cultural model can be contrasted with an important middle-class Japanese cultural model that stresses an individual's harmonious fit with his or her social environment. Culture scientists regard these cultural models as key ingredients in the overall emotional experience.

Although these scientists rely on cultural models as a context for understanding and predicting emotional responses, they do not assume that all people within a culture are alike. Because individuals in a culture will regard a given model in different ways, their individual emotional experiences will differ as well. For example, not all Americans value personal achievement equally, and they certainly differ in how they define achievement personally or in how they go about accomplishing success. What is critical, however, is that the overall culture still powerfully endorses the dominant cultural beliefs and social practices for how things are done in that culture, for example, when we are encouraged to "just do it" by a sports equipment company. The cultural models, therefore, set the reality boundaries within which emotions occur and are expressed. According to culture researchers, the models that are important to a culture will influence several aspects of emotions.

3.1. Antecedent Events

Cultural models are important in determining the occurrence of the events that cause an emotion to occur. How an event transpires is not part of the emotion itself, but it is an important determinant of which emotion will occur and what the responses of the individual will be. In other words, certain situations may be especially likely to lead to certain emotions. If these situations occur more often in one culture than in another, the accompanying emotion is also more likely to occur in the former culture than in the latter. Likewise, if a situation is relatively unlikely to occur in a particular culture, the emotions that typically go with that situation are probably less likely to occur as well. How likely an event is to occur depends, at least in part, on how well it agrees with the overall values of the culture. Cultures tend to create more opportunities for events that are consistent with the prevailing models, whereas they tend to suppress opportunities for events that are at odds with these models. For instance, consistent with the cultural models, Americans create and encourage many contexts in which people are made to feel good about themselves because they stand out (e.g., praise, compliments, awards, trophies). The creation of many of those events has been shown to result in high occurrences of happy feelings. Similarly, in cultures that value a harmonious fit with other people (e.g., Utku Inuit, Chewong of aboriginal Malaysia, Tahitians, Japanese), acts of upsetting and frustrating those around a person are largely absent. The absence of frustrations is reflected in the relatively low levels of anger reported in these cultures.

People also selectively pay attention to certain aspects of situations in ways that are consistent with the cultural models. People in cultures that stress autonomy and uniqueness (e.g., United States) focus on good events and good qualities in the self so as to feel unique and independent. Therefore, individuals in this kind of culture are predisposed to strive for positive rewards. This focus on approaching positive outcomes is known as a promotion focus. In cultures where harmonious interaction with others is emphasized (e.g., Japan), individuals strive to avoid the occurrence of negative events by examining what could be improved about oneself. This focus on the prevention of bad outcomes (e.g., not living up to the standards) is referred to as a prevention focus. These cultural differences in predispositions for attending to certain events have implications for the emotions experienced as well. Consistent with their prevention focus, Chinese individuals report relaxation or relief when goals have been met and report anxiety when goals have not been reached. Americans' promotion focus leads them to report happiness when goals have been met and to report feelings of sadness when goals have not been reached. Thus, cultural models form the emotional lives of individuals within the culture by promoting events that are consistent with the models

and by focusing people's attention on particular aspects of situations that occur.

3.2. Appraisal

Cultural models also influence the ways in which people appraise or evaluate events. Certain appraisals are more likely to occur in certain cultures than in others. The prevalence of a particular appraisal will also differ across cultures, depending on the various cultural models. For instance, the appraisal of agency (responsibility) is markedly different across cultures. A key aspect of North American cultural models is success through independent personal accomplishment. Claiming responsibility and a personal sense of control are at the center of what it is to be a good person in this culture. In keeping with this model, North Americans typically attribute personal "responsibility" (either to oneself or to others) in many situations (especially positive ones), and they strive for a state of responsibility and control.

In contrast, East Asian cultural models emphasize the role of fate, how events are determined by many factors other than individuals themselves, and how individuals are interwoven with others in their social environment. Individual control and responsibility are not very relevant to the East Asian way of thinking. Accordingly, East Asians do not view control (or lack thereof) as an important aspect of emotional situations.

Thus, agency appraisals are markedly different across cultures. These evaluations of personal control or responsibility have important roles in the emotional experience. One implication of feeling responsibility, or feeling that someone else has responsibility, is that a person can influence or control other people in specific ways that fate or the world in general cannot. Claiming responsibility for an event results in active ways of emotionally reacting to, and coping with, the event. This differs considerably from the case of an event where one believes that the event could not have been humanly controlled.

Cultural models can also account for differences in the most frequent patterns of appraisal. For example, research has shown that there are differences in the events that are associated with feelings of pleasantness. In cultures that stress individual autonomy, pleasant experiences are those that are associated with evaluations of high self-esteem and personal control, whereas in cultures that stress interpersonal engagement and harmony, pleasantness is associated with events that make one feel connected with other people. Thus, pleasantness can be a different experience in different cultures, depending on what each culture associates with pleasantness.

3.3. Relationship Goals and Action Readiness

Cultural models also outline desired goals for relationships with other people, thereby affecting an individual's impulses in a situation. In a series of studies in which people from various cultures reported what they felt like doing in a variety of emotion situations, Americans reported that they desired to (a) affirm or salvage positive self-esteem, (b) actively change the situations to make them consistent with their personal goals, and (c) prioritize their own personal concerns over those of their relationships. For example, in response to negative interactions, Americans reported the urge to break off their relationships and protect self-interests at the expense of their relationships. In contrast, Japanese reported action readiness modes that served to (a) maintain or improve their relationships and (b) adjust to other people's wishes and expectations. For instance, in negative interactions with others, Japanese wanted to reestablish closeness with the other persons and to keep their emotions to themselves.

Cultural differences in emotions are not random. Rather, they make sense when one considers the role of cultural models. The influence of cultural models goes beyond providing the content to fill in the blank of an otherwise standard emotion process in that cultural models affect the entire process of emotional experience. The context in which an emotion occurs, the ways in which the situation is perceived and evaluated, and the coping that is generated by the emotions experienced all are guided by cultural models. Thus, the cultural approach focuses on cultural differences instead of aiming to establish universalities. It is important to study both similarities and differences for a complete body of knowledge about emotions. However, the claim that emotions are primarily universal has been challenged by cultural researchers. During recent years, the methods adopted by cultural studies have established many cultural differences in emotion.

4. CONCLUSION

So, can one accurately interpret the emotions of people from other cultures? Few researchers of emotion are likely to answer with an unconditional "yes." However,

scientists' answers differ according to the ways in which they approach the question. Researchers who seek basic emotions believe that the most important emotions are similar across cultures. Researchers who have analyzed the separate building blocks of emotions in search of similarities contend that emotions in other cultures can be understood if one knows how certain situations are appraised and reacted to in other cultures. Finally, culture researchers believe that one must know and understand the major cultural concerns, meanings, and relational practices to have a sense of the emotional experiences in other cultures.

See Also the Following Articles

Emotion ∎ Intelligence, Emotional ∎ Motives and Goals

Further Reading

Kitayama, S., & Markus, H. R. (Eds.). (1994). *Emotion and culture: Empirical studies of mutual influence.* Washington, DC: American Psychological Association.

Levenson, R. W., Ekman, P., Heider, K., & Friesen, W. V. (1992). Emotion and autonomic nervous system: Activity in the Minangkabu of West Sumatra. *Journal of Personality and Social Psychology, 62,* 972–988.

Mesquita, B. (2003). Emotions as dynamic cultural phenomena. In R. Davidson, H. Goldsmith, & P. Rozin (Eds.), *Handbook of affective sciences* (pp. 871–890). New York: Oxford University Press.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin, 112,* 179–204.

Shweder, R. A., & Haidt, J. (2000). The cultural psychology of the emotions: Ancient and new. In M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 397–414). New York: Guilford.

Emotion in Sport: An Individualized Approach

Yuri L. Hanin

KIHU – Research Institute for Olympic Sports, Jyvaskyla, Finland

GLOSSARY

experience The totality of characteristics, both past and present, that make up the particular quality of a person's performance; three aspects of experience are situational experience (states), relatively stable patterns of experience, and meta-experience (knowledge, attitudes, and preferences for [or rejections of] a person's experiences).

individual zones of optimal functioning An action-oriented framework for analysis of idiosyncratic emotional experiences related to athletic performance.

in–out of the zone principle The assumption that specific intensity of emotion has a beneficial (or detrimental) effect on individual performance; a high probability of individually successful performance is related to being in the optimal zone of intensity.

psychobiosocial state A situational, multimodal, and dynamic manifestation of the total human functioning.

resources Psychobiosocial (cognitive, affective, motivational, bodily, behavioral, operational, and communicative) assets determining an athlete's ability to perform up to his or her potential.

Understanding emotion–performance relationships is of crucial importance for applied sport psychologists. This article reviews an individual- and action-oriented framework, termed the Individual Zones of Optimal Functioning (IZOF) model, that describes, predicts, and explains individually optimal and dysfunctional emotional experiences related to athletic performances. To describe emotion–performance relationships, the IZOF model proposes a multidimensional conceptualization of emotions as a component of the psychobiosocial state. Finally, the model suggests several empirically substantiated guidelines for person- and task-relevant assessments focused on emotion self-regulation and optimization of performance.

1. INTRODUCTION

One of the key questions in sport psychology is how and why pleasant and unpleasant emotional states affect

sport performance. Indeed, are unpleasant emotions always bad for performance? For instance, can anger sometimes be useful in competition? On the other hand, are pleasant emotions always good for performance? For instance, can self-confidence lead to poor performance through complacency and underestimation of task demands? Moreover, do athletes perform up to their potential when they are stress free or when they are deliberately using competitive stress to their advantage? Is precompetition anxiety bad for all athletes? Finally, is it possible to predict an athlete's performance based on content and intensity of emotional experiences? Researchers, applied sport psychologists, athletes, and coaches often need straightforward answers to these practical questions. Most of the existing models in sport psychology focus mainly on interindividual differences between successful and unsuccessful athletes. On the one hand, there are exceptional athletes who reach the top and stay at the top with remarkable consistency. On the other hand, it is important to realize that even the winners have bad days and often exhibit inconsistent performance. In other words, performers across the full range of skill and experience are quite familiar with occasional performance slumps and "choking." Thus, instead of comparing high achievers with underachievers, an alternative approach would be to focus more on performance experiences of each individual athlete. Such an approach has direct practical implications and can be applied to a particular athlete or a team. This article describes an individual-oriented, research-based framework, termed the Individual Zones of Optimal Functioning (IZOF) model, to show how the understanding of emotion–performance relationships at the individual level can be helpful in optimization of athletic performance.

2. INDIVIDUALIZED APPROACH TO PERFORMANCE EMOTIONS

Successful performance enhancement interventions in sport should (a) focus on unique challenges and concerns facing coaches and athletes, (b) emphasize practical application of research-based principles, and (c) realize that athletes are healthy achievement-oriented people who are "forced" to cope with performance-induced stress. The IZOF model aims to match these requirements with a special emphasis on individual athletes. The following subsections briefly describe the basic assumptions of the model.

2.1. The Meaning of the IZOF Acronym

The IZOF model, developed in high-achievement sport settings during the past 30 years or so by Hanin, is a sports-specific, action-oriented framework. It focuses on analysis of the structure, dynamics, and function of situational emotional experiences (e.g., nervous, worried, calm, confident) related to individually successful and less than successful performances. The model also examines relatively stable patterns of these experiences and meta-experiences (i.e., beliefs about expected effects of feeling nervous, worried, or calm on performance). The IZOF acronym reflects selected features of the model.

2.1.1. Individual

The model is an idiographic approach that focuses on idiosyncratic emotional experiences of individual athletes (or teams). Therefore, the main emphasis is on the within-individual dynamics of subjective emotional experiences and meta-experiences related to performance.

2.1.2. Zones

The zone principle implies a specific relationship between the perceived intensity of emotional state and the quality of performance. Individually optimal intensity (high, moderate, or low) reflects an athlete's past performance history and available resources. Being in the optimal zone implies a high probability of individually successful performance. Thus, the zones serve as empirically established individualized criteria for prediction of performance, and the "in–out of the zone" concept is used to evaluate a match (or mismatch) between the currently experienced level of emotion intensity and previously established optimal and dysfunctional zones.

2.1.3. Optimal

Optimal emotions are defined as most relevant and appropriate for a particular athlete performing a specific task. The optimal performance state results in a total task involvement and the best recruitment and use of available resources. This usually results in a high-quality performance process and achievement of individually successful performance outcomes (although not necessarily "peak or best ever"). The optimal performance state tends to include both pleasant and unpleasant emotions reflecting idiosyncratic strategies that athletes use in recruitment and use of their resources. Therefore,

the functionally optimal state (e.g., situational anger, high anxiety) is different from the "ideal" or preferred state (e.g., "flow"), which usually has only positive characterisms. Finally, the concept of optimality is multidimensional and is related not only to emotion intensity and content but also to time and context dimensions.

2.1.4. Functioning

The term "functioning" emphasizes the fact that a detailed description of performance emotions is just a tool to understand their optimal and dysfunctional effects on task involvement and the quality of performance process. Optimal function is manifested in an athlete's efficient recruitment (effort) and use (skill) of available resources. This usually results in energizing and organizing effects on performance. In contrast, emotion dysfunction usually reflects a failure to recruit resources and their inefficient use, resulting in deenergizing and disorganizing effects of emotion on athletic performance.

2.2. Basic Assumptions of the IZOF Model

The IZOF model holds that emotion is a component of adaptive or maladaptive response in person–environment interactions; emotion is an integral part of a person's psychobiosocial state, performance process, and total functioning; emotion is triggered by appraisals of a person's relationships with the environment; and emotion is a reflection of critical moments of imbalance in person–environment interactions. The model describes the nature of the emotion–performance relationships as follows. First, emotions are triggered by a person's appraisal of the probability of achieving expected individually relevant goals. Second, because sport activity is repetitive, situational emotional experiences gradually develop into relatively stable emotion patterns. Third, emotion patterns are specific to the individual, task, and setting, and this specificity is manifested in the form, content, intensity, time, and context dimensions of the psychobiosocial state. Fourth, emotion–performance relationships are reciprocal. Emotions affect performance; however, changes in the ongoing performance process and intermediate (and final) outcomes produce a shift in a personal meaning of the situation, resulting in a shift in emotion content and intensity. A shift in the direction of impact (causality) is usually determined by ongoing appraisals of the performance process. Fifth, prediction of performance should be based on

assessment of interaction effects of a cluster of really experienced emotions (optimal and dysfunctional) rather than on separate effects of selected "pure" emotions. Sixth, because athletes often reflect on their experiences in successful and less than successful performances, they gradually develop meta-experiences. For instance, an athlete could notice that high anxiety sometimes helps him or her to focus better or that it does not affect performance negatively. This knowledge leads to a positive attitude toward anxiety, which is then interpreted as an indicator of this athlete's readiness for competition. Emotion self-regulation and individual-oriented interventions in the IZOF model are based on the concept of optimal personal meaning and adequate meta-experiences triggering specific emotions.

3. MULTIDIMENSIONAL DESCRIPTION OF PERFORMANCE EMOTIONS

The IZOF model defines the performance-related psychobiosocial state as a situational, multimodal, and dynamic manifestation of the total human functioning. A multilevel and system description of performance-related states includes at least five interrelated dimensions: form, content, intensity, time, and context. Three of these dimensions (form, content, and intensity) describe the structure of subjective experiences and meta-experiences, and two of these dimensions (time and context) characterize the dynamics of performers' subjective experiences. When viewed collectively, these five basic dimensions provide a framework for a relatively complete description of the phenomenon and can guide data collection, analysis, interpretation, and generation of new ideas.

3.1. Form Dimension

An athlete's psychobiosocial state manifests itself in seven basic forms (modalities): cognitive, affective, motivational, bodily, motor–behavioral, operational, and communicative. All seven modalities are interrelated and provide a relatively complete description of a performance state, including experiences and their displays (expression or suppression). Thus, emotion is construed as an important component of the performance-related psychobiosocial state. Current individual-oriented research focuses on affective (emotional), motivational, and bodily components of the performance state and their interactive effects.

3.2. Content Dimension

The content dimension is a qualitative characteristic of the performance state. The IZOF model as an alternative individual-oriented approach categorizes emotion content within the framework of two closely related but independent factors: a hedonic tone (pleasure–displeasure) and functionality (optimal–dysfunctional impact of emotion on athletic performance). Both factors reflect qualitatively different experiences related to individually successful and poor performances. The four broad emotion categories derived from these distinctions are pleasant and functionally optimal emotions (P+), unpleasant and functionally optimal emotions (N+), pleasant and dysfunctional emotions (P–), and unpleasant and dysfunctional emotions (N–). These four categories provide an initial robust and sufficiently broad structure to generate a wide range of idiosyncratic, individually relevant, and task-specific emotions really experienced by athletes prior to, during, and after successful and less than successful performances. Finally, the content of these athlete-generated emotions can be recategorized to see how idiosyncratic descriptors match or mismatch the existing "basic" or discrete emotion frameworks proposed in mainstream psychology.

3.3. Intensity Dimension

The intensity dimension as a quantitative attribute of subjective experiences is expressed in either objective or subjective metrics on a selected component or modality of the performance state. Different levels of intensity can produce different functional effect related to energizing (deenergizing) and organizing (disorganizing) aspects of the performance process. The in–out of the zone notion describes a selected range of emotion intensity producing optimal, neutral, or dysfunctional effects on individual performance. For instance, a high probability of successful performance is expected if an athlete's actual state is in (or close to) the earlier established optimal zone of intensity. In contrast, if an athlete's actual state is out of the optimal zone, the athlete is likely to perform below his or her potential.

3.4. Context Dimension

The context dimension is an environmental characteristic reflecting the impact of situational, interpersonal, intragroup, and organizational factors on emotion intensity and content in sport setting. Situational impact is observed in emotional reactions triggered in practices and competitions by athletes' anticipated or real contacts and interactions with significant others (e.g., partners, coaches, teammates). This dimension also includes culturally coded and culturally determined beliefs of participants about the expected impact of specific emotions on their performance and about the rules of emotion display in a particular subculture.

3.5. Time Dimension

The time dimension includes topological (e.g., phases, cycles, sequencing, periodicity, timing) and metric (e.g., duration, frequency) characteristics reflecting the dynamics of performance-related experiences. The short-term dynamics in sporting activity is observed in emotional states prior to preparation for an action, during task execution, and after performance in a single competition (practice). The long-term temporal dynamics is related to emotion–performance relationships during a competitive season(s), a 4-year Olympic cycle, or an athlete's sports career.

4. PREDICTION OF EMOTION–PERFORMANCE RELATIONSHIPS

There are certain myths in sport psychology about emotion–performance relationships based on unsubstantiated beliefs that have been challenged recently. Some examples include the notions that there exists a universally optimal intensity of emotions facilitating performance of all athletes (e.g., moderate anxiety), that all negative emotions are always dysfunctional for athletic performance, and that all positive emotions are always optimal for athletic performance. In contrast, IZOF-based individual-oriented and empirically supported predictions of emotion–performance relationships indicate that these notions are not the case.

4.1. Interindividual Variability

As predicted, research revealed a high degree of interindividual variability in the intensity and content of idiosyncratic optimal and dysfunctional emotions accompanying individually successful and poor performances. Thus, different athletes can perform up to their potential experiencing emotions of different content and intensity. Optimal emotional states reflect (a) individual differences in available resources, (b) athletes' ability to recruit and use these resources

efficiently, and (c) athletes' individual coping strategies to compensate for either a lack or an insufficiency of resources (low situational readiness for competition).

4.2. In–Out of the Zone Principle

The prediction of individually successful, average, or poor performance based on the in–out of the zone principle has received some empirical support in several studies. Thus, a player's emotional state is first described in terms of individually relevant emotion content and intensity. Then, current emotion intensity is contrasted with the previously established individually optimal and dysfunctional intensity zones. A high probability of successful performance is expected when emotion intensity is within the optimal zones and outside dysfunctional ranges. Therefore, a large discrepancy (in intensity and emotion content) between the actual state and the individually established optimal zones indicates a high probability of less than successful performance.

4.3. Interactive Effects

Research indicates that most emotional experiences involve clusters of "mixed" rather than "pure" emotions. For instance, an international-level tennis player, prior to his best ever match, felt high intensity of pleasant optimal emotions (P+). He felt highly determined, confident, excited, dynamic, and comfortable. At the same time, he felt moderately aggressive, alarmed, and somewhat uncertain (N+). Moreover, his unpleasant dysfunctional emotions (N–) (e.g., nervous, afraid, worried, intense) were of low intensity. This pattern was similar to that during his best match; however, he felt alert and quick but not excited too much without premature satisfaction. In contrast, in the worst match of his career, this player felt highly nervous and worried (N–), and these experiences were more intense during the match. Interestingly, at the same time, his optimal pleasant emotions prior to and during the match were of moderate and low intensity, respectively. If applied sport psychologists measured only anxiety, for instance, they would miss the entire picture of emotional clusters. Therefore, the interactive effects of emotions enhancing and impairing sporting activity are of crucial importance for an accurate prediction of emotion–performance relationships. In this case, a high probability of individually successful performance is expected when combined maximum-enhancing and minimum-impairing effects are observed. On the other hand, a high probability of individually average and impaired performance is expected when a combination of high-enhancing and high-impairing effects or of low-enhancing and low-inhibitory effects is observed. Finally, a high probability of poor performance is expected when low-enhancing and high-inhibitory effects are observed.

4.4. Dynamics and Bidirectionality

Emotion–performance relationships are dynamic and bidirectional. In other words, precompetition emotions affect performance, and ongoing performance affects the dynamics of mid- and postcompetition emotions. Thus, to predict the entire process of emotion–performance relationships, it is important to establish patterns of emotion impact on performance and of performance impact on emotions.

All of these predictions were tested empirically across various samples and sports in different countries. Initially, the IZOF model focused on precompetition anxiety as a discrete stress-related emotion syndrome with a "fixed" emotion content. The main emphasis in the IZOF–anxiety research was on identifying individually salient intensity of state anxiety. Specifically, it was revealed that each athlete had an individually optimal intensity level (which could be high, moderate, or low) and zones of anxiety enhancing the athlete's performance. The analysis of an athlete's past performance history identified these optimal level and intensity zones, which served as tentative individualized criteria in the prediction of current and future performances. Thus, the probability of successful performance was high when current precompetition anxiety was near or within the previously established individually optimal intensity zones. When precompetition anxiety fell outside the zones (i.e., higher or lower), individual performance usually deteriorated. Recently, similar results were obtained in the studies of optimal and dysfunctional effects of situational anger on athletic performance.

Research indicated that an exclusive focus on precompetition anxiety or anger was limited because these syndromes were only a part of athletes' emotional experiences related to performance. Therefore, the IZOF model was extended to the study of pleasant and unpleasant emotions enhancing and impairing individual performance, and more consideration was given to the saliency of other dimensions of emotion (content, time, and context). A recent extension of this line of research tested the validity of the preceding predictions in motivational, bodily, and performance correlates of the psychobiosocial state.

5. EXPLAINING EMOTION–PERFORMANCE RELATIONSHIPS

Most of the existing approaches in sport psychology focused mainly on antecedents and consequences of selected emotion syndromes (e.g., anxiety, anger, depression). In contrast, the main emphasis in the IZOF model was first on an accurate and detailed description of various forms of idiosyncratic emotional experiences in athletes. To explain the antecedents (emergence and dynamics) and consequences (impact on performance) of performance-related emotional experiences, the model uses several constructs that are briefly reviewed in the following subsections.

5.1. Appraisals and the Emotion Process

Person–environment interactions are important in the IZOF model; therefore, a psychosocial perspective is central in the functional interpretation of the dynamics of emotion–performance relationships. First, emotion is conceptualized as an unfolding process reflecting person–environment interactions. Ongoing appraisals of these interactions result in a change in personal meaning of the situation, and this exerts the influence on emotional experiences related to performance. Second, changes in personal meaning reflecting the dynamics of the performance process trigger functional shifts in emotion content and intensity. Specifically, anecdotal evidence and practical experience indicate that functionally optimal pleasant and unpleasant emotions (P+N+) prior to and during activity are usually anticipatory and are triggered by the appraisals of challenge and threat. In contrast, situationally dysfunctional pleasant and unpleasant emotions (P–N–) during activity are usually triggered by perception of "achieved" outcomes (appraisals of gain and loss) before the task is completed.

5.2. Emotion Impact on Performance

Two constructs with their opposites related to energizing and organizing aspects of emotion account for the impact of emotions on the athletic performance process. These include the concepts of energy mobilization (and demobilization) and energy use (and misuse). Therefore, optimal and dysfunctional emotion function in the IZOF model is conceptualized within the framework of two closely related but independent factors: (a) energy mobilization (e.g., optimal effort, intensity) and (b) energy use (e.g., efficiency, optimal information processing). The former is related to the situational resources available to an individual performer, whereas the latter characterizes the efficiency of using these resources. Based on these two factors, four relatively independent global emotion functions are derived: (a) energizing or energy-mobilizing function, (b) energy demobilizing function, (c) energy use or regulation function, and (d) energy misuse or deregulation function. These four functions provide a framework for interpretation of separate and interactive effects of positive and negative emotions on individual performance. Based on the nature of these interactions, the total impact of emotions on athletic performance can be optimal (regarding effort and skill), paraoptimal (with only effort or skill, but not both, being optimal), or dysfunctional (regarding effort and skill).

From the functional effect perspective, the constructs of energy mobilization and use (and their opposites) seem to be useful in explaining why in some athletes optimal emotions are predominantly pleasant, whereas in other athletes they are unpleasant. For instance, low-anxious athletes are typically "smart users" of available energy and are less distracted by task-irrelevant and energy-wasting concerns. In contrast, high-anxious athletes typically generate more energy, especially in stressful or emergency situations, because they are often less efficient in its use due to a narrow attention focus and an overload in the information-processing function. Thus, unpleasant emotions, such as anxiety, are functionally useful for these athletes in that they help to generate additional energy to compensate for the apparent limitation in information processing or the use of energy. Because the quality of athletic performance is usually related to both the amount of available energy and its efficient use, different athletes can be successful by using different resources. In other words, the same level of performance may be achieved either through the increase of total effort or through skillful (smart) use of available resources (efficiency).

6. INDIVIDUALIZED ASSESSMENTS AND MONITORING OF EMOTIONS

An accurate individualized assessment and monitoring of performance-related emotional experiences provides a basis for predictions of emotion impact and self-regulation. In the following subsections, selected data collection techniques that have been used to assess

individually relevant idiosyncratic emotion content in sport are briefly described: interviews, emotion profiling, self-generated metaphors, and narratives.

6.1. Interviews

The first and most popular data collection technique for eliciting idiosyncratic emotion content is structured and semistructured in-depth interviews with open-ended and probing questions activating an athlete's recall of thoughts and feelings prior to and during best and worst performances. It is usually recommended in pilot and exploratory studies to generate idiosyncratic labels and then to aggregate most selected items into a stimulus list or standardized emotion scale. Identifying idiosyncratic content and personally meaningful labels to describe athletes' subjective experiences is a clear advantage of these interviews over questionnaires with researcher-generated items and "fixed" emotion content. Experienced and verbally skillful athletes can provide detailed and meaningful accounts of their experiences (how they felt) and meta-experiences (how they interpreted and coped with these feelings) prior to and/or during the competition. Athletes who are less aware of their performance-related experiences might require some assistance initially in structuring and focusing their recall.

6.2. Individualized Emotion Profiling

A person's perspective is central in the reconstruction of athletes' emotional experiences and meta-experiences related to successful and less than successful performances. The individualized emotion profiling aims to analyze athletes' past performance history and significant emotional experiences and meta-experiences. Specifically, athletes generate individually relevant emotion words that best describe their optimal (helpful) and dysfunctional (harmful) pleasant and unpleasant emotional experiences accompanying their performances in best and worst competitions. A stimulus list, compiled from existing emotion scales or based on previously generated items, can assist athletes in generating emotion descriptors. The IZOF-based emotion list consists of 46 pleasant emotions (arranged in 14 synonymous rows) and 50 unpleasant emotions (arranged in 14 synonymous rows). Table I shows a group-oriented functional categorization of pleasant and unpleasant emotions based on several studies of athletes representing various sports. However, research also indicates that in each individual case, optimal and dysfunctional emotions can be different from those in the aggregated group data. Therefore, individualized emotion profiling is strongly recommended.

TABLE I
Functionally Categorized Emotion Stimulus List for Individualized Profiling

Pleasant emotions	Unpleasant emotions
P+	N+
• Active, dynamic, energetic, vigorous	• Angry, aggressive, furious, violent
• Quick, rapid, fast, alert	• Intense, fierce
• Confident, certain, sure	• Tense, strained, tight, rigid
• Determined, set, settled, resolute	• Annoyed, irritated, distressed
• Brave, bold, daring, dashing	• Anxious, apprehensive, worried
• Inspired, motivated, stimulated	• Concerned, alarmed, disturbed, dissatisfied
• Excited, thrilled	• Nervous, jittery, uneasy, restless
P–	N–
• Delighted, overjoyed, exhilarated	• Discouraged, dispirited, depressed
• Relaxed, comfortable, easygoing	• Afraid, fearful, scared, panicky
• Calm, peaceful, unhurried, quiet	• Doubtful, uncertain, indecisive, irresolute
• Cheerful, merry, happy	• Helpless, unsafe, insecure
• Glad, pleased, satisfied, contented	• Inactive, sluggish, lazy
• Lighthearted, carefree	• Sorry, unhappy, regretful, sad, cheerless
• Nice, pleasant, agreeable	• Tired, weary, exhausted, worn out
Your own emotion:	Your own emotion: . . . .

Individualized emotion profiling involves several steps. First, to identify optimal emotion patterns, athletes select four or five positive items and then four or five negative items that best describe their emotions related to their individually successful performances in the past. Then, athletes identify their dysfunctional emotions related to individually poor performances by selecting four or five positive items and four or five negative items. Although athletes use the emotion stimulus list to generate individually relevant descriptors, they can also add emotion words of their own choice. The focus of recall is either on a specific situation (precompetition) or on repeated experiences across several similar situations. Individually optimal and dysfunctional intensities for each emotion item are established using the Borg's CR-10 scale (ranging from "nothing at all" to "maximal possible").

In sum, individually oriented and task-specific emotion profiling identifies (a) individually relevant emotion content, (b) individually optimal and dysfunctional emotion intensity, (c) individually relevant context, and (d) interaction effects of optimal and dysfunctional emotions. In a comprehensive IZOF-based assessment program, first, the individual's past performance history is examined within the context of the current situation. Then, individually relevant emotion content is generated. Next, optimal (and dysfunctional) intensities for each emotion item are identified. Finally, the emotion profiles (with intensity zones for each individual item) are validated and refined in a series of repeated self-ratings in specific performance situations (practice or competitions). For monitoring emotional states in a field setting, the total emotion profile can be condensed to a few most important markers (e.g., three items for each of the four emotion categories). The major emphasis in individualized emotion profiling is on analysis and reconstruction of athletes' performance-related idiosyncratic experiences and meta-experiences by using their own vocabularies.

Research shows that athlete-generated emotion descriptors usually have strong connotations with nonemotion concomitants of the psychobiosocial state such as cognitive (e.g., alert, focused), motivational (e.g., willing, determined, motivated), and bodily (e.g., relaxed, tense, jittery) experiences. Therefore, there is a clear need to extend the focus of individualized profiling from emotions to other modalities of an athlete's state.

Moreover, a mere description of emotional experience (e.g., the same high level of precompetition anxiety in an athlete on two occasions) might not explain why the athlete was very successful in the first competition and less than successful in the second competition. Here, additional qualitative data (e.g.,

motivational domains or causes of these experiences and resulting self-talk) could be instrumental. For instance, a highly anxious athlete in a successful race could feel "psyched up," "perceive the race as important," and "focus on competing with tough opponents." This athlete's self-talk would be clearly self-empowering (e.g., "I can run at this level," "I'll show them all," "It is my day today"). In contrast, prior to a less than successful race, this athlete's thoughts included "doubts in her current shape," "insufficient recovery," "health problems," and/or "poor team climate," and the resulting self-talk was clearly self-defeating in that there was too much focus on opponents' strengths (e.g., "they are going to beat me," "I've got no chance in this race," "today is not my day," "the race is meaningless").

6.3. Self-Generated Metaphors

Although the individualized emotion profiling may be successful in generating idiosyncratic descriptors of emotions, there is still a need for a more holistic description of the psychological meaning of the performance situation and an athlete's role in it. To address this concern, Hanin and colleagues proposed a metaphor self-generation method to examine the feasibility of symbolic representation of performance-related experiences using idiosyncratic athlete-generated metaphors. The approach is based on the symbolizing ability of people and on the notion that people give structure, meaning, and continuity to their activities by symbolizing their experiences.

A metaphor is always related to a certain image, which is holistic and meaningful for a person. Therefore, words, objects, mental images, stories, and the like are the units of metaphor. A symbolic nature of a metaphor provides a picture of one's understanding of something, with an emphasis on most personally relevant and important aspects and shades of meaning. Well-known metaphors describing performance-related states in sport include being "in the zone," "in the groove," "in the cocoon," and "in the flow state" as well as having "choked." Research suggests that athlete-generated idiosyncratic metaphors may be useful in the practice of sport psychology for holistic and personally meaningful description and better understanding of performance-related situational experiences and relatively stable patterns of experiences.

The metaphor self-generation method is an instrument identifying metaphors and accompanying interpretive descriptors of feeling states prior to, during, and after best and worst competitions. Thus, the

focus is on assessment and symbolic representation of idiosyncratic content (quality) of emotional experiences in a high-achievement setting. Several studies using a metaphor generation method in various performance situations revealed that athletes in various sports, and at various age and skill levels, are able to generate highly idiosyncratic images symbolically representing their high or low action readiness in successful and poor performance situations.

Metaphors reflecting available resources (e.g., "a boat with an engine," "a mountain river") and ability to recruit resources (e.g., "a horse in a light cart," "a man who can move a mountain") have been used to describe athletes' states in best ever competitions. A lack of resources (e.g., "a sinking boat," "an empty bottle") and inability to recruit resources (e.g., "a man with weights on his legs," "a bird unable to fly") are characteristic of metaphors that athletes have used to describe their states in worst ever competitions. In addition, metaphors describing best ever competitions have reflected highly efficient use of resources (e.g., "Batman in flight," "a fisherman catching a golden fish"), whereas metaphors describing worst ever competitions have reflected inefficient or poor use of resources (e.g., "a bear after hibernation," "a soldier in a burning tank").

An example of using metaphors to monitor feeling states of a Finnish Olympic-level track and field male athlete in a series of international competitions during the season illustrates a practical utility of the method. A chain of metaphor images generated by this athlete for various contests included "a colt on a spring pasture" (very strong but too excited to control his own performance), "a tractor" (started moving in the right direction), "a Volvo" (good routines, self-confident, and reliable performance), "a Mercedes" (ready and knew what to do, enjoying successful performance), and "a Ferrari" (feeling great, powerful, and in control). These metaphors helped to identify and anchor the athlete's optimal level of confidence, the right focus, and self-control skills prior to and during competition.

Taken together, research findings indicate that idiosyncratic content and personally relevant meaning of metaphors are either self-empowering (usually in best competitions) or self-defeating (usually in worst competitions) images. Furthermore, the quality of metaphors is culturally coded and determined by the specific contexts in which athletes perceive and experience performance-related episodes. Context-induced beliefs about potential effects of pleasant and unpleasant emotions can have a strong impact on perception of consequences of emotional experiences for athletic performance. Thus,

athletes can perceive competitive stress either as an undesirable factor to be avoided or as a part of their job with which they must cope. Consequently, if athletes believe that high precompetition anxiety is detrimental to their performance, the quality of such anxiety will be different from that of anxiety perceived as an indicator of high readiness for the competition. Therefore, before attempting any intervention program, sport psychologists must identify these context-determined beliefs based on past idiosyncratic experiences.

Research findings also indicate that metaphors differentiate various contexts (e.g., best ever and worst ever competitions) and capture the dynamics of personal meaning across pre-, mid-, and postcompetition situations. For instance, in best ever competitions, metaphors symbolizing strength, power, speed, and skillfulness were generated 5 to 10 times more often than in worst ever competitions. In contrast, in worst ever competitions, metaphors symbolizing weakness, slowness, and powerlessness were generated 5 times more often than in best ever competitions. Similar and even stronger relationships were observed regarding the efficient recruitment and use of available resources.

Idiosyncratic metaphors symbolically representing the content of feeling states experienced throughout competitions were different (low content overlap) across pre-, mid-, and postcompetition situations. These data provide support for the notion that a performance task usually involves three interrelated, but functionally and qualitatively different, stages: preparation for, execution of, and evaluation of performance. The findings also suggest that the functional and personal meaning of these three situations, as represented symbolically by idiosyncratic metaphors, is different for each athlete. A change in meaning triggers a change in emotional states. In other words, emotion dynamics is observed not only in the intensity of emotions but also in the content (quality) of emotions. Therefore, in research and applications, it is important to identify a specific constellation (a set or cluster) of emotion content that is optimal or dysfunctional for an athlete's performance. Moreover, metaphors reflecting even slight changes in personal meaning of each performance situation could be useful markers for applied sport psychologists in monitoring the dynamics of performance states and in evaluating the effectiveness of individualized interventions.

6.4. Narratives

A narrative is an account of several interconnected occurrences. It is a scheme by means of which people

give meaning to their experiences. It also provides a framework for understanding the past events of one's life and for planning future actions. It is the primary scheme by means of which human existence is rendered meaningful. As Bruner argued in 1991, narratives are self-stories conveying intention and feeling related to a sequence of events; therefore, they provide a holistic picture of unfolding processes. Narratives (as self-stories told by athletes and coaches) can describe concrete performance situations and can identify automatic thoughts and emotional responses. Thus, they help to recreate a high-level performer's experiences, meta-experiences, and available (or unavailable) coping strategies.

The following transcript illustrates how a 16-year-old ice hockey goaltender perceives and describes his problem during important games:

> This is perhaps a silly question, but can someone help me? I am very good technically and physically with the qualities needed for a goaltender. I always train much and well. For instance, I trained on the average 12 times per week during the summer season. However, I have one very big problem. In the game, I start feeling very nervous and fear that I will play badly. Then usually easy things become very difficult. . . . This is especially often in the games where it is necessary to win. Sometimes the tension goes away if the play starts well. I know that I am really good. My coach thinks that my flexibility and other qualities are top level in Finland. . . . In training, I can tend off really well, but in really tough situations, I cannot stand it. . . . Now it is again the time when I have to succeed, especially in training games. Based on these results, they will select a future national junior team.

This narrative describes briefly the goaltender's feelings and performance in practices and tough games, accompanied by his general self-ratings and expectations. However, the main advantage of narratives is that they provide a means for holistic and dynamic reconstruction of the unfolding process of athletic performance in a real-life setting. Narratives focus on the functions of emotion feelings and displays in the episodes of everyday life (or a specific activity). Thus, of special value is the fact that storytelling includes not only emotion words as descriptors of real idiosyncratic experiences but also an athlete's interpretations of these experiences and culturally determined beliefs within a specific context (i.e., meta-experiences). Although the main emphasis so far was on emotion experiences, it is important to realize that emotion displays (i.e., expression and suppression of feelings)

also have specific features and stereotyped views in the context of competitive and high-achievement sport. In this setting, athletes also learn to read the body signals of their opponents to identify the opponents' feeling states and action readiness.

7. PRACTICAL IMPLICATIONS

Applied sport psychologists assisting athletes in assessments and self-regulation of emotions usually face three issues: (a) identification of emotional states related to individually successful and poor performances, (b) prediction of emotion–performance relationships, and (c) selection of person- and task-relevant techniques of self-regulation. Compelling empirical evidence in support of the IZOF model provides several tentative guidelines on how to deal with these three issues.

First, to identify individually optimal and dysfunctional emotional experiences, establishing an individually relevant cluster (constellation) of emotions and their intensities prior to, during, and after successful and less than successful (poor, average, or customary) performances is recommended. These qualitatively and quantitatively extreme situational experiences serve as individualized criteria in the evaluation of current and anticipated emotional states. In addition, it is important to identify athletes' specific beliefs and attitudes about their emotion impact on performance. Are they aware of such effects? How do they usually cope with stress- and complacency-producing situations? Are these situational emotional experiences random or relatively stable patterns that athletes try to reproduce in important competitions? The main purpose of such individualized assessments is to enhance athletes' awareness and acceptance of these experiences.

Second, prediction of emotion–performance relationships is based on the in–out of the zone notion using categorical or continuous approaches. A categorical approach predicts performance based on the comparison between previously established individual zones and actual scores of intensity. A continuous approach is based on perceived "intensity impact" contingencies along the entire working intensity range of each emotion. Here, the emphasis is on an estimation of partial and total effects rather than on only a selected optimal range of each emotion. In both cases, a decision about emotion regulation (increase or decrease) is based on the magnitude of deviations either from optimal and dysfunctional zones or from a total effect in the selected modality. Furthermore, the intervention should aim not

Conventional and IZOF-Based Features of Emotion Assessment and Self-Regulation

Basic principle	Conventional feature	IZOF-based feature
Multimodality	One to three components	Seven basic forms
Multirange/-content	Fixed/Same for all	Idiosyncratic
Multizone/-intensity	Moderate or nonspecified	IZOF/Idiosyncratic intensity
Multistage	Mainly precompetition	Pre-, mid-, or postcompetition
Multidirection	Decreases or nonspecified	Individualized up/down
Multitask	Nonspecified	Specified and identified
Multifunction	Nonspecified	Effort or skill focused
Multimethod	One or two researcher generated	"Own," tailored or combined

only at helping athletes to enter or reenter their optimal zones but also at staying away from the dysfunctional zones. Finally, predictions should also consider the total functional impact of emotion on performance that is usually manifested in an increase or a decrease in effort (energy) and in efficiency or inefficiency in the use of available resources.

Third, there are numerous techniques of emotion self-regulation in the practice of applied sport psychology. Effective emotion self-regulation should be based on already described individualized assessments and predictions of emotion–performance relationships. Moreover, a selected method or intervention strategy (technique) should be optimal for both the athlete and the situation. In other words, the method should match previously established individual patterns of coping with emotion-triggering situations. In addition, the effective intervention program usually includes several appropriate methods of self-regulation.

Finally, there are several directions for effective emotion regulation. Most of them focus directly on emotional response by using various mental skills. However, there are other options that include a change of the current situation or its perception (personal meaning) by an athlete or a special organization of athletic activity for an athlete or a team (role expectations and game tactics). Table II shows IZOF-based and conventional (group-oriented) features reflecting several practical guidelines that are important to consider in assessments and regulation of emotional states.

8. CONCLUSION

The main purpose of this article was to provide an overview of the IZOF model as an alternative

individual-oriented idiographic approach to studying performance-related emotions. The IZOF model suggests some tentative answers to the questions asked at the beginning of the article.

First, unpleasant emotions as strong situational experiences are not always detrimental to athletic performance. Thus, anxiety, anger, and/or tension could be helpful in generating more energy and effort, and sometimes they could even compensate for a situational lack of needed resources. These emotions, if well channeled on task process, can postpone fatigue, sustain alertness, and keep the right focus. Coaches and athletes, as well as sport psychologists, begin to realize that coping with stress and using stress to enhance performance are possible.

Second, pleasant emotions are not always beneficial for performance, especially in sports requiring sustained focus, effort, and persistency for a relatively long time. A special problem in a high-achievement setting often presents a demotivational impact of excessive complacency following repeated success. Moreover, high self-confidence can sometimes lead to complacency or the underestimation of a "weak" opponent, resulting in insufficient alertness, lack of focus, and/or carelessness. These, in turn, can have a debilitating effect on performance, often leading to unexpected and season-ending injuries. In such cases, idiosyncratic experience is the best indicator of how athletes can perform up to their potential, either stress free or using competitive stress to their advantage.

The IZOF model provides a conceptual framework and methodological tools to describe, predict, assess, and explain situational emotional experiences related to performance, relatively stable patterns of these experiences, and meta-experiences. It is a clearly idiographic and person-oriented approach, with the main focus on

TABLE III
Comparison of Group-Oriented and IZOF-Based Models

Feature	Group oriented	IZOF based
Analysis level	Global/Contextual	Situational
Target	Groups, samples	Individual athletes
Emphasis	Between-individuals	Within-individual
Content	Global/Narrow	Task and person relevant
Performance	Unrelated	Related to, induced by
Setting	Nonspecific	Sport specific
Time	Static, stable	Dynamic, changeable
Assessment	Group referenced	Self-referenced
Scales	Standardized	Individualized

within-individual comparisons that provide the basis for athlete-tailored and interindividual generalizations. Although the IZOF model is consistent with several existing orientations in psychology, it differs mainly in its emphasis on intraindividual dynamics in emotion–performance relationships (Table III). From the applied perspective, the major advantage of the individualized approach to studying emotion–performance relationships is its ability to capture and explain the unique findings that are often missed or ignored in group-oriented models. However, both individualized and group-oriented perspectives can be useful in applied sport psychology.

See Also the Following Articles

Anxiety and Optimal Athletic Performance ■ Arousal in Sport ■ Assessment in Sport Psychology ■ Attention and Concentration Training in Sport ■ Cheating in Sport ■ Competition in Sport ■ Decision Making in Sport ■ Emotion ■ Fair Treatment and Discrimination in Sport ■ Goal Setting and Achievement Motivation in Sport ■ Group Dynamics in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Performance Slumps in Sport: Prevention and Coping ■ Psychological Skills Training in Sport ■ Self-Confidence in Athletes ■ Sport Psychology, Overview ■ Successful Athletic Careers

Further Reading

Hanin, Y. L. (Ed.). (2000). *Emotions in sport*. Champaign, IL: Human Kinetics Press.

Hanin, Y. L. (2002). Individually optimal recovery in sports: An application of the IZOF model. In M. Kellmann (Ed.), *Enhancing recovery: Preventing underperformance in athletes* (pp. 199–217). Champaign, IL: Human Kinetics Press.

Hanin, Y. L. (2003, February). Performance-related emotional states in sport: A qualitative analysis. *Forum: Qualitative Social Research*, *4*(1). Available: www.qualitative-research.net/fqs-texte/1-03/1-03hanin-e.htm.

Hanin, Y. L., & Stambulova, N. B. (2002). Metaphoric description of performance states: An application of the IZOF model. *The Sport Psychologist*, *16*, 396–415.

Robazza, C., Pellizzari, M., & Hanin, Y. (in press). Emotion self-regulation and athletic performance: An application of the IZOF model. *Psychology of Sport and Exercise*.

Ruiz, M. C., & Hanin, Y. L. (in press). Metaphoric description and individualized emotion profiling of performance related states in high-level karate athletes. *Journal of Applied Sport Psychology*, *16*.

Employment Discrimination

Mary L. Tenopyr

Bridgewater, New Jersey, USA

GLOSSARY

actuarial prediction Predicting future behavior only on the basis on actual data; very much like predictions made in the insurance industry.

Age Discrimination in Employment Act of 1967 (ADEA) A U.S. federal law that prohibits discrimination against persons over 40 years of age.

Americans with Disabilities Act of 1990 (ADA) A broad federal law that prohibits discrimination against persons with disabilities in many areas, including employment.

bias A systematic effect; that is, the same degree of bias occurs for every member of the group of interest.

Civil Rights Act of 1964 A comprehensive equal opportunity federal law that includes prohibition of employment discrimination on the basis of race, color, religion, or national origin.

concurrent study A criterion-related study of employment methods that involves obtaining predictor and criterion data at approximately the same time.

criterion-related validity The relationship between test results and a criterion that is a measure of an outcome of legitimate interest to an employer.

differential item functioning (DIF) The possibility that some items on an employment test function differently for distinct groups; it is difficult to conduct proper studies in this area, but it has become apparent that DIF is not found very often.

disparate impact theory The requirement that an employer justify the employment method, even if it seems neutral, if the employment method has a disparate impact on a legally defined group.

disparate treatment theory Intentional discrimination not necessarily involving tests or employment methods.

Equal Employment Opportunity Commission (EEOC) A U.S. enforcement agency created by the Civil Rights Act of 1964.

fairness A social term; there are various conceptions of what is fair given that what appears to be fair to one person might not be seen as fair by another person.

flagging the score A practice (considered controversial) of communicating to various members of management whether an employment test was administered under specially accommodating conditions.

holistic assessment Essentially a method for predicting future behavior on the basis of judgment given assorted data of several kinds; it is akin to clinical prediction by psychologists.

internal structure Any of the relationships among items or scales comprising a test; it can affect reliability or replicability of test results.

predictive bias An effect that occurs when a group of interest does significantly better or worse than would be predicted from scores for the whole group of applicants; the groups

of interest in this situation are differentiated on some characteristics not relevant to criterion performance.

predictive study A criterion-related study of employment methods that yields an indication of how accurately test results can predict criterion scores that are obtained at a later time.

race norming Selecting employees in terms of their standing in their own group and not on their standing in the whole job applicant population.

test Any procedure used in making employment decisions; this broad definition is frequently used in professional standards and guidelines, and it reflects the fact that employment discrimination is possible over a wide range of employment methods and not just paper-and-pencil testing.

validity The accuracy of inferences drawn from test results.

validity generalization The extent to which one can generalize from results of analyses and apply them in a particular employment context; with due care, a professional wishing to use tests can review the published validity generalization evidence and make an informed decision about what selection procedures to use in a given situation.

The selection of employees, whether in terms of hire, promotion, or other personnel actions, is a matter of concern relative to civil rights. Issues regarding the application of psychology in making selection decisions have become a major focus of legal activity. Employment discrimination on the basis of membership in some specific groups is prohibited by law. Legal challenges are frequently directed toward the methods used in employee selection. In particular, certain types of psychological tests and their use have come under close legal scrutiny. Concepts that have been associated with testing, such as validity and reliability, have been incorporated into various government guidelines and regulations that have, in turn, greatly influenced the development of laws relating to civil rights. Testing concepts and their application have been subject to considerable debate, research, and (in some cases) redefinition. Bias in selection is now better understood, and associated concepts have been evaluated thoroughly. It is clear that there are different psychological considerations associated with the selection and treatment of different groups. Despite advances in knowledge relative to employee selection, there are still research needs related to selection of some groups. Nevertheless, the large amount of psychological study that has been devoted to understanding employment discrimination and its many aspects has added greatly to the database of psychological knowledge.

1. BASIC CONCEPTS IN EMPLOYMENT DISCRIMINATION

There are many situations in which there are fewer desirable rewards than there are seekers of those rewards. For example, except in unusual circumstances, there are fewer job openings than there are candidates for the jobs. Thus, there must be a means of selecting from among the job seekers. An analogy to a gatekeeping function may be drawn; in this case, the employment process serves as a gate, and the gatekeeping involves the application of psychology. In particular, the psychology of individual differences and its associated procedures have been widely applied in making employment decisions.

No two adult individuals are exactly alike, and each can be expected to behave in unique ways in many situations. Even identical twins, depending on the circumstances, can show noticeable differences in behavior. Within every definable group in society, there are individual differences, and often there is a wide distribution of these differences in, for example, scores on a test.

In addition, relative to some psychological processes and achievements, there are consistent differences among groups on average. Many of the gatekeeping functions in employment involve procedures that indicate average differences among groups. The term "average" is important because it represents a situation far less dramatic in terms of discrimination than many believe. For example, because of individual differences, there will generally be a number of different scores within each definable group. The distributions of scores for the separate groups will probably overlap. That is, given two groups, some members of the group that has lower scores on average will score higher than the group that has higher scores on average. Obviously, when there are average differences among groups on a test or another selection procedure, there is the potential for discrimination on the basis of group membership. Discrimination based on group has undoubtedly been a feature of every employing society in the world. So long as there are discernable group differences— and there always will be some differences with regard to some aspects of the employment process—there is the possibility of discrimination.

2. EARLY LAW RELATING TO EMPLOYMENT DISCRIMINATION

In many ways, the laws regarding employment discrimination started with a relatively narrow focus in the United States. However, during recent years, new legal developments have resulted in the scope of the issues being increased, and new aspects of psychology have come to the forefront. For example, new research has been done relative to the employment of the aging and of persons with disabilities, and the results of that research have been widely applied.

The abolishment of slavery by the 13th Amendment to the U.S. Constitution in 1865 and the Reconstruction Era after the Civil War brought forth the first legal approaches to employment discrimination. Thus, race was the focus of early laws in this area. Unfortunately, equal treatment of Black Americans in most aspects of life really did not receive much impetus until the 20th century. Significant early events of this era were President Franklin Roosevelt's 1941 order to end discrimination in the defense industry, President Harry Truman's 1948 order to desegregate the military services, and President Dwight Eisenhower's 1957 order to send troops to enforce desegregation of a high school in Arkansas. Thus, the stage was set for the emergence of the modern civil rights movement that involved a major march on the U.S. capital in 1963, numerous other marches and demonstrations, and a series of major urban confrontations.

These events greatly influenced congressional action on the Civil Rights Act of 1964, a comprehensive equal opportunity statute that included prohibition of employment discrimination on the basis of race, color, religion, or national origin. Many of the aspects of this broad law and subsequent amendments and court interpretations have influenced the employment practices of all major employers into the 21st century. The 1964 act had profound effects on the direction and focus of applied psychology in the workplace. This came about mainly due to an amendment prompted by a then recent situation in Illinois. A Black job applicant had filed a complaint with the state's fair employment commission because he was denied employment as a result of failing a cognitive ability test. A commissioner ordered the test replaced by a more culturally balanced one. This case received considerable national attention and was the impetus for considerable congressional debate, which ended in the eventual passage of an amendment that, on the surface, appeared to protect employers' use of ability tests in employment.

This amendment stated that it is not unlawful for an employer to give or act on the results of any professional developed ability test, provided that the test is not designed, intended, or used to discriminate. Mainly as a result of this amendment and other developments in the civil rights area, there was an extraordinary growth of interest in differences in group cognitive ability, especially in applied areas. Academic journals were flooded with research articles, and various theories and mathematical concepts of discrimination were debated. Furthermore, new studies related to group differences were extended beyond those relating to cognitive abilities. For example, as women moved into physically demanding jobs, studies of sex differences in various job-related physical aspects multiplied. Because the Civil Rights Act of 1964 specifically designated testing as a matter of concern, various agencies at both the federal and local levels began to promulgate guidelines on employee selection. All of these emphasized testing, and all other employee selection procedures, such as the interview, were defined as tests. The position taken by the U.S. Equal Employment Opportunity Commission (EEOC), an enforcement agency created by the 1964 act, was that a "professionally developed ability test" is one that is "professionally developed and used." This definition, which implied that validity (i.e., the accuracy of inferences drawn from test results) is the primary basis for professional use, appeared in the brief EEOC guidelines issued in 1964. In 1970, the agency issued detailed prescriptions for studying and evaluating validity. Furthermore, professional organizations began to issue their own guidelines on employee selection, and various professional standards on testing were modified.

3. ROLE OF VALIDITY

Validity, its definition, methods for conducting validation studies, and methods for assessing possible discrimination have become increasingly complex over the years. Testing standards have been revised not only to respond to technical advancements but also to address new issues associated with testing of groups that became the focus of civil rights acts passed after 1964. Nevertheless, validity remains a unifying theme in many of the laws involving possible discrimination in employment. Many local laws, which are even broader in group coverage than are federal laws, incorporate validity concepts.

The emphasis on validity in the 1970 EEOC guidelines was codified through the landmark 1971 U.S. Supreme Court decision in *Griggs v. Duke Power*

Company. In this case, the court ruled that if a test tends to exclude Blacks, the employer has the obligation to demonstrate that the test is manifestly related to the requirements of the job. Thus, business necessity could justify employment procedures. Also, the court addressed the question of relative qualifications and how high employment standards could be set. The novel feature of the *Griggs* decision was the introduction of disparate impact theory. Under this theory, if an employment method, although neutral of its face, has a disparate impact on a legally defined group, the employer must justify the employment method. Intention to discriminate is not a necessary element of a disparate impact case. Intentional discrimination is covered under a separate theory called disparate treatment theory. The emphasis on validity was continued in various government pronouncements at both federal and local levels. The most influential of these was the Uniform Guidelines on Employee Selection Procedures issued by several federal agencies in 1978.

The complicated further development of case law and legislation in the area of employment discrimination resulted in change to the conduct of a case brought under disparate impact theory. The Civil Rights Act of 1991 delineated three steps in a disparate impact case. First, the plaintiff must prove that the use of the selection procedure has a disparate impact on his or her legally defined group. The burden of proof then shifts to the defendant, who must prove that the procedure is job related (i.e., valid). The plaintiff then has the option of showing that a specific equally valid selection procedure not having the same degree of adverse impact could have been used. The 1991 act incorporated case law that eliminated an onerous provision of the 1970 EEOC guidelines requiring that an employer using a selection procedure that had a disparate impact must prove that there was no suitable alternative. In formal logic, this is known as "proving the negative"—which cannot be done.

4. DISCRIMINATION IN EDUCATION

Constitutional rights apply to both the education and employment venues. However, many of the laws relating to discrimination in the field of education are not the same as those governing employment. Nevertheless, many concepts developed in law relevant to education have been applied in employment contexts. For example, "diversity" as a suitable employment goal evolved

from a U.S. Supreme Court decision in 1978, when the issue was admission to medical school. In a subsequent major and narrowly tailored decision on university admissions in 2003, the court allowed race and other factors to be taken into account in university admissions. The reason why the court allowed for this action is to achieve student body diversity. This is to be done on a time-limited basis, and no fixed quotas may be set. In addition, weighting techniques, such as point systems that favor some groups over others, are prohibited. The court supported a holistic system involving a number of methods such as testing, essay writing, and interviewing. This action brings into focus the major psychological question of the validity of clinical prediction as opposed to actuarial prediction. The holistic assessment endorsed by the court is essentially a method for predicting future behavior on the basis of judgment given assorted data of several kinds and is akin to clinical prediction. Actuarial prediction, very much like predictions made in the insurance industry, involves predicting future behavior only on the basis on actual data. For example, the quality of work that would be done by a job applicant, if hired, is predicted from actual test scores.

5. VALIDATION

In an employment situation, the inferences made through the application of employee selection procedures are predictive in nature regardless of the nature and format of the methods used in making the personnel decisions. Throughout this discussion, the term "test" is used to cover any procedure used in making employment decisions. This broad definition of test is frequently used in professional standards and guidelines and reflects the fact that employment discrimination is possible over a wide range of employment methods and not just paper-and-pencil testing.

The term "validity" has had a number of definitions over the years. Today, validity is not seen as a property of the test. Instead, it refers to the accuracy of inferences made on the basis of test scores. For example, if it is inferred that persons with low test scores will not do the job as well as will persons with higher test scores, and this inference is confirmed as a result of proper study, one may say that the inference is accurate and, hence, valid.

Validity is not an all-or-none concept; there are varying degrees of validity. Validation is the process of conducting a study to assess the degree of validity. In this case, the term "validation" is not used in its other sense as an approval process.

The legal use of the term "job related," as opposed to "valid," clearly and appropriately narrows the scope of any validation evidence to that related to the workplace activities of the concerned persons. It should be noted that activities other than actual job performance can be the focus of the test. For example, tenure, training grades, accidents, and absences may be legitimate concerns of the employer, who may use tests designed to predict them.

Furthermore, it is important to recognize that validity in psychology, as in many fields, is never perfect. There are always errors in prediction. However, the lack of perfect prediction need not necessarily deter an employer from using a test. How useful a test will be depends very much on the particular situation. Under some conditions, a test with relatively low validity can be quite useful to the employer. Certain factors, such as labor market conditions and job difficulty, must be taken into account. Professional judgment, as well as available data, should inform decisions on test use.

6. EVIDENCE OF VALIDITY

There are five generally recognized sources of evidence. Sources for evidence of validity are based on (a) content, (b) response processes, (c) internal structure of the test, (d) relationships of test scores with other variables, and (e) consequences of test use.

6.1. Evidence Based on Content

Content validation involves an initial thorough analysis of the work in terms of work behaviors and activities and/or knowledge, skills, abilities, or other personal characteristics related to job demands. The work analysis may be as detailed as need be, depending on the situation. Again, professional judgment is required. A work analysis for a parachute packer job should be rigorously detailed. For other jobs that are less rigid in terms of critical requirements, a broader analysis might be appropriate.

A selection procedure developed on the basis of content need not necessarily cover the whole job. If it is critical that every task involved in a job be done precisely, every task should be covered in minute detail. However, for many jobs, only some of the job duties are of critical interest, and only these parts may be covered by the test. For example, a person doing word processing may also answer the telephone on occasion. The employer may test on word processing skills and not test on telephone answering skills. Measuring abilities of various sorts presents particular problems. In general, content validation is appropriate only for relatively narrow abilities. If one wants to measure broad abilities such as leadership, content should not be the primary source of evidence.

Developing a test on the basis of content is essentially a sampling process in which a sample of the work or the worker characteristics of interest is taken. Evaluation of the adequacy of the sampling is difficult. Validity is established mainly on scrutiny of the processes used in sampling from the content area of interest and the procedures used in developing the test. Professional judgment is usually the only feasible way in which to assess the adequacy of the evidence. However, some psychologists contend that evidence can be gathered through judgments regarding job tasks or the skills, knowledge, abilities, and other characteristics required for the job and their linkage to the content of the test.

Many of the disputes involving employment discrimination have arisen in situations involving validity evidence based on content. Because of the large judgmental component involved, there is opportunity for both employment discrimination and disputes about the adequacy of the sampling process used and the product developed. Also, many of the tests used in the public sector are based on content, and much of the case law regarding employment discrimination has been generated in the public sector. In particular, selection for public safety jobs that are sought by thousands of applicants, such as police work and firefighting, is a source of contention. Most public sector merit systems have legal rules and rigid requirements for employee selection. For example, job applicants are put in rank order in terms of standing on tests, and the final selection must be from among the top three candidates. Furthermore, lists of candidates may be used to fill openings for months or even years, and applicants might have a long wait before a new test is developed and they have the opportunity to be tested again. Private employers generally have more options in the use of selection procedures and, consequently, can make adjustments to the selection process when there are questions about employment discrimination.

There are other issues related to applied psychology in the use of content-based evidence. One of these is whether it is appropriate to rank order candidates on the basis of scores on a content-based test. It appears that most industrial psychologists believe that such ranking is permissible. Another related issue is how to set "cut" (i.e., passing) scores on such tests. There

are still disagreements about how high passing scores should be set. The main issue is whether they should be consistent with the average proficiency of the workforce or whether an employer trying to improve the company's workforce may set higher standards. Another is whether changing test formats, such as moving from a paper-and-pencil test to a more complicated and costly procedure involving pictures or exercises that simulate job duties, will reduce disparate impact on various groups. The Civil Rights Act of 1991 allows cost to be a consideration in determining whether specific alternative selection procedures should be adopted.

A specific source of contention is the application of content-based physical ability tests to both genders. The differences in performance between the genders on many physical ability tests is marked on average, and issues related to ranking, passing scores, and the possibility of alternative selection procedures becomes magnified. For example, a content-based test for firefighters might involve activities such as dragging a firehose up a flight of stairs. There have been questions about whether simpler physical tests, including those involving activities such as push-ups, can be justified on the basis of content. Persons with certain types of physical disabilities have raised similar questions about such tests.

Some employers have representatives of various groups that are likely to be subject to discrimination review the test content to detect possible discriminatory items or those that might be offensive to particular groups. There are questions about the effectiveness of this procedure, and more research concerning the value of this method in preventing discrimination is needed.

6.2. Evidence Based on Response Processes

Where employment is concerned, evidence based on response processes (e.g., how long an examinee requires to respond to a particular type of item), is not so relevant as most of the other types of evidence of validity. Response processes do have some value in determining how a test is functioning and what abilities the applicant is using in taking some types of tests.

These processes have been assessed by various methods such as having an examinee talk about what mental techniques he or she is using while trying to solve an arithmetic problem. Physical measurements based on recording eye movement are often used, as are observations by judges. It has been suggested that

such monitoring of response production may eventually be able to discern differences in the ways in which different groups respond to the same test material. However, relevant literature on research in employment settings is lacking.

6.3. Evidence Based on Internal Structure

The third type of evidence is based on the internal structure of the test. An example of internal structure would be the extent to which test items are in the same format such as solving analogies problems. Internal structure can affect reliability or replicability of test results. Such reliability is important in supporting validity. In general, validity is limited by lack of reliability of test results. Reliability may be estimated by various methods. The choice of exact estimation procedure depends on a number of factors. In some respects, a finding of validity is, to some extent, evidence of reliability.

Relative to discrimination, the possibility that some items function differently for distinct groups is often studied. This effect is called differential item functioning (DIF). It is difficult to conduct proper studies in this area, but it has become apparent that DIF is not found very often. When found, reasons for such differential functioning are not usually easily identified.

6.4. Evidence Based on Relationships with Other Variables

The fourth type of evidence, relationships of test scores with other variables, has been the major source of statistically based definitions of discrimination and of much confirmable information in psychology about group differences. This type of evidence will undoubtedly be a mainstay in studying employment discrimination and related topics for some time to come.

In an employment situation, the relationship of most interest is the relationship between test results and a criterion that is a measure of an outcome of legitimate interest to the employer. This relationship is called criterion-related validity, and the generic predictive model of validity is merely a generalization of the criterion-related model. Despite the fact that criteria can take many forms, supervisors' ratings appear to be the most commonly used criteria. The design of a criterion-related study can involve a number of specific procedures, but these can be subsumed under the general headings of "predictive" and "concurrent" studies.

A predictive study yields an indication of how accurately test results can predict criterion scores that are obtained at a later time. For example, job applicants may be tested and the results might not be used in hiring. Then, at a later date, job performance or other job outcome data that form the criterion may be collected. A concurrent study involves obtaining predictor and criterion data at approximately the same time. Typically, present employees are used in these studies. In the area of employment, it is very difficult to conduct predictive studies and complete them in a reasonable period of time. Consequently, much of the validation relevant to employee selection is based on concurrent designs. For cognitive ability, it appears that validation results from predictive and concurrent studies do not vary much from each other. More studies need to be done relative to other types of tests.

Relative to discrimination, one of the major findings resulting from the application of the criterion-related model is that validities in the cognitive ability area do not differ greatly between Caucasians and African Americans. As the United States becomes more multi-cultural, there is a need for research with newly emergent racial and cultural groups. Also, it is necessary to determine the extent of group differences relative to tests that reflect characteristics other than cognitive ability.

Differences between gender groups are a matter of concern. However, they are difficult to study because during test development, particularly of the most widely published cognitive ability tests, efforts have been made to reduce some gender differences. Still, there are some well-documented findings relative to possible discrimination. Women are at a disadvantage when selection procedures involve some spatial and mechanical abilities. Also with respect to cognitive ability tests, women generally have a narrower range of scores than do men. Some of the practical effects of this phenomenon are that there are more males than females classified as mentally retarded, and there are more males than females who qualify for prestigious scholarships and other honors that are awarded on the basis of high-level cognitive ability.

6.5. Evidence Based on Consequences

Within a broad social context, concern about the consequences of testing, particularly that some types of testing can result in unequal distribution of rewards to various groups, has led some major researchers to define validity partially in terms of consequences of test use. In an employment context, consequences may influence decisions regarding the use of particular selection practices. However, when validity is defined as it has been historically, consequences alone do not form acceptable evidence of validity. Socially undesirable consequences in employment should prompt a thorough investigation of possible sources of problems and corrective actions available.

6.6. Other Evidence-Related Issues

With the possibility of pressure to substitute clinical prediction for actuarial prediction in employment operations, there is a real need to conduct further studies on the decision processes involved in clinical judgments as they relate to employment. Questions about validity and effects on groups are a matter of concern. Past findings clearly demonstrate that the actuarial prediction process produces valid results to the same extent as does clinical prediction. In many cases, actuarial prediction is associated with greater validity. Among the major problems that have plagued clinical prediction in the past is the basing of judgments on methods that have not been shown to produce valid results. Despite the fact that use of multiple methods forms the basis for clinical prediction, the possibility that some of these methods do not add to the accuracy of such prediction should be considered. This is particularly important when tests that are not designed for use in employment inform the clinical judgment process used in selecting new hires from among job candidates. Another problem is that judges might not weigh the available information optimally. Too much weight can be attached to the results of methods that yield marginal validity.

An extraordinary amount of evidence has been accumulated regarding the validity that can be expected from the application of various employee selection methods. A mathematical technique known as meta-analysis has afforded the opportunity to analyze the results of accumulated research so that general conclusions relative to various research findings can be drawn. In employment psychology, meta-analysis has been applied to the accumulated validity studies done over a period of many decades. The extent to which one can generalize from results of these analyses and apply them in a particular employment context is known as "validity generalization." With due care, a professional wishing to use tests can review the published validity generalization evidence and make an

informed decision about what selection procedures to use in a given situation. The published validity generalization research results have been based on a substantial amount of data in some areas, including those involved with cognitive ability; however, in some other areas, the evidence is still accumulating. Nevertheless, it is clear that most professionals consider the validity generalization data first before they make decisions about the use of selection procedures.

7. FAIRNESS AND BIAS

Although the term "fairness" was used in early federal and local guidelines on testing, it has become customary in applied psychology to distinguish between "fairness" and "bias." The former is a social term; what appears to be fair to one person might not be seen as fair by another person. Thus, there are various conceptions of what is fair. Conversely, latter term has a specific meaning in employment psychology. This definition is based on results of research with the criterion-related model and is generally accepted in applied psychology.

First, fairness may be viewed as equitable treatment in the employment process. For example, all applicants should have the same selection procedures applied to them under the same conditions. Second, fairness may be conceptualized as all applicants having the same opportunity to learn material covered by employment tests. With a few exceptions, this notion, conceived in the context of educational achievement testing, does not apply to employee selection. Third, fairness may be viewed in terms of whether the employment process provides equal outcomes for different groups. In other words, the same proportion of each group involved in the process should be selected. In the interest of affirmative action, various techniques have been applied in attempts to achieve equal outcomes. One method is to add points to the scores of members of selected groups. Another method is race norming, that is, selecting employees in terms of their standing in their own group and not on their standing in the whole job applicant population. In this case, a person from a group that has low average scores would be more likely to be selected than would a person with the same test score if the latter person belongs to a group with higher average scores. Other mathematically complicated methods have been used with the objective of reducing group differences in selection. For example, a method known as the sliding band, which seeks to lessen the disparate impact associated with ranking systems in the public

sector, has been developed. Considerable controversy has arisen regarding the application of these and other methods developed in the context of affirmative action. Fourth, fairness may be seen as a lack of "predictive bias," which is defined technically as occurring when a group of interest does significantly better or worse than would be predicted from scores for the whole group of applicants. The groups of interest in this situation are differentiated on some characteristics not relevant to criterion performance. Examples of such characteristics might be race and gender. Bias is a systematic effect; that is, the same degree of bias occurs for every member of the group of interest. So far as African Americans and Caucasians are concerned, when cognitive ability tests are used in employment, bias is seldom found. This situation might not be seen when other groups are compared. The rapidly changing composition of the labor force, coupled with the need for large samples of research participants, renders the proper conduct of studies of possible bias very difficult except in military organizations and other large employers.

A number of other issues are related to more general questions of fairness and bias. It is well known that many groups are judged in terms of stereotypes. Gender, ethnic, racial, and other stereotypes can affect employment decisions. Considerable research in the area of social psychology has been devoted to stereotypical thinking as it applies to employment. Some controversy still prevails regarding the application of this research.

As a result of women being paid less than men, there are both federal and (some) local laws requiring equal pay for equal work that involves the same duties. To determine whether the work done is indeed equal requires the work analysis methods developed by applied psychologists. Legal issues concerning pay equity for the genders continue. There have been major efforts to extend the laws to require equal pay for jobs of comparable worth. The determination of comparable worth for jobs with different job duties is extremely challenging for industrial psychologists.

8. SPECIAL SELECTION SITUATIONS

8.1. Aging

Employment issues involving possible age discrimination are covered by special laws and regulations both at the federal level and in many local jurisdictions. The Age Discrimination in Employment Act of 1967 (ADEA) prohibits discrimination against persons over

40 years of age. As the population ages and people extend their working lives past what was once considered normal retirement age, issues relevant to employment discrimination against the aging have increased both in scope and in intensity. The use of measures of cognitive ability is a particularly sensitive issue. It is now known that many of the major abilities decline with age more slowly than had once been believed. However, the possible onset of debilitating diseases, such as Alzheimer's disease, must be considered in some cases. Also, there may be interactions among mental functioning, physical condition, and other variables (e.g., gender).

Test results might not have the same meaning for the aging as they do for other groups protected by law. At the ages when individuals over 40 years old are likely to seek employment, major age-related deficits in mental functioning are relatively unlikely. When there are deficits, the more likely case is that the individuals involved have had long histories of obtaining low scores on measures relating to cognitive functioning. Another possibility to be considered is that discrimination against older persons may be associated with other characteristics such as gender and ethnic origins.

It appears that it would be unrealistic to strictly apply all of the law that has been developed under the Civil Rights Act of 1964 and its progeny to the aging population. The application of disparate impact theory is particularly controversial. The issues relative to mental functioning of older persons are more complicated than most of those involving younger people.

Considerable legal attention has been directed toward the various processes involved in terminating the employment of older persons. Any techniques used in making termination decisions, such as performance appraisals, should be treated as tests and subjected to the same professional scrutiny. The economic issues involved in retirement and other forms of termination have been the focus of considerable legal attention. Among the major situations that involve the potential for age discrimination is the provision of health care benefits to retirees. An understanding of the needs for mental health care for the elderly can inform employer decisions regarding health care benefits for both retirees and older persons still employed. Psychologists can provide useful input in this area.

8.2. Individuals with Disabilities

The Americans with Disabilities Act of 1990 (ADA) has received considerable attention by applied psychologists because the statute applies to both mental and physical disabilities. This broad law prohibits discrimination against persons with disabilities in many areas, including employment. This law involves a number of ambiguities and, thus, is subject to continuing legal interpretations. Some of the requirements of the law remain to be clarified. It appears that the courts are progressively limiting the scope of ADA. Furthermore, solid research bases to support actions under this statute are unavailable. To conduct proper research in this area, one needs a substantial number of research participants with the same disability and the same level of symptoms. Such large groups of people are generally unavailable. Basically, ADA recognizes the individual nature of disabilities and the need to treat each person with a disability in a different way. ADA is basically an extension of earlier acts dealing with rehabilitation and employment of veterans.

Under ADA, disability is defined as a physical or mental impairment that substantially limits at least one major life activity. According to EEOC guidance, among the activities covered are learning, working, thinking, concentrating, and interacting with others. Specifically excluded are certain conditions such as current illegal drug use, a number of conditions dealing with sexual behavior and gender identification, certain compulsive disorders such as kleptomania, and impairments that are expected to be temporary. In addition, employers are not expected to hire individuals who may be a danger to themselves or others. Questions about diagnoses clearly involve the application of psychological knowledge about mental impairments. In addition, making predictions about future behavior is necessary.

The Americans with Disabilities Act of 1990 applies only to qualified persons with disabilities. These individuals must be able, with or without reasonable accommodations, to perform the essential functions of the job. Determining the essential functions of the job may involve an adaptation of the work analysis methods used in conventional employee selection situations. Such analysis is not legally required, but many employers conduct it. During prehire questioning, such as might be done on an application blank or during an interview, the employer cannot ask questions that would require the applicant to indicate the existence of a disability. However, the employer can give a psychological examination unless the test is used for medical reasons, in other words, to determine whether the individual is mentally impaired or otherwise mentally disabled. Diagnostic psychological tests and related procedures may be used in a medical

examination, which can be given only after a job offer has been made, with the same questions being asked of all applicants for the same job category.

In testing situations before a job offer is made, the job candidate may ask for an accommodation relative to the medium of testing, test format, testing conditions, or testing time. For example, a person with impaired vision may ask for a test booklet with enlarged print. The most commonly requested accommodation is extra time to take the test. Accommodations are to be done on an individual basis. Nevertheless, psychologists try to be reasonably consistent in devising accommodations for similar types of impairments. In many cases, psychologists collaborate with physicians and other professionals in developing accommodations in testing and other aspects of the employment process.

Most employers ask job candidates for documentation regarding disabilities that are not obvious when either tests or job accommodations are requested. Documentation regarding mental disabilities is usually carefully scrutinized. Many employers have internal requirements for documentation. For example, a family practice physician's diagnosis of a learning disability might not be accepted. Accommodations in working conditions on the job are often developed with input from several types of specialists. Some employers have standing committees to make decisions about job accommodations. Most job accommodations are not particularly expensive. In any case, the employers are not expected to undergo undue hardship. In general, legal interpretations of what constitutes hardship have been quite narrow.

The issue of indicating whether a test was administered under accommodating conditions should be recorded, and whether this information should be communicated to various members of management is quite controversial. This is called "flagging the score." At a minimum, a psychologist usually makes some record of the accommodation and maintains it in a filing system that is separate from other test score files. Various other legal and practical issues are related to the employment of persons with disabilities. In particular, violence in the workplace might be related to the failure of the employer to recognize that a person claiming a disability was a danger to himself or herself and others. Other aspects of workplace security, such as developing a crisis management plan in the event of disaster (natural or otherwise), may involve paying special attention to the needs of persons with disabilities. Of obvious importance are plans for the evacuation of such individuals. Psychologists can provide input to a number of other aspects of such plans and can advise on training and practice needs.

The role of disabilities with respect to a number of other laws may require the attention of psychologists. Included among such laws are the federal Family and Medical Leave Act and the various workers' compensation and unemployment compensation laws, all of which may have some features that can lead to discrimination against persons with mental disabilities.

8.3. Linguistic Minorities

Discrimination on the basis of national origin may occur due to the language in which a test is written or administered. Linguistic considerations in employment are not straightforward. One of the first distinctions to be made is that between a language and a dialect. Some have contended that tests should reflect dialects shared by sizable groups of people. As the United States becomes more multicultural, attending to all of the various dialects, whether regional or group associated, would be an onerous task for even the largest employers. If job proficiency in the language of the workplace is required, employment processes usually use the language needed to communicate on the job. A distinction should be made between oral communication and written communication. For example, a given individual's lifestyle may involve proficiency in writing but not in speaking. This could occur in cases where English was the language of the schools attended, but some other language was spoken in the home. Others may speak in a mixture of two languages, and others with backgrounds in two languages may have difficulty in using both. Any translations of tests should be done by persons qualified not only in the languages but also in psychometrics. There are a number of procedures necessary to ensure that the original test and the translated version yield equivalent results.

9. SEXUAL HARASSMENT

A major issue in the area of employment discrimination is sexual harassment. This subject has been a major source of litigation since the early 1970s. Sexual harassment can be physical, verbal, or visual. It is not always directed toward women, but in any event it is always unwelcome. Legally, the conduct should be, at a minimum, pervasive or severe enough to create a hostile or abusive work environment and be

such that a reasonable person would find it to be offensive. In legal terms, sexual harassment unreasonably interferes with work performance, is physically threatening or humiliating, and involves more than isolated utterances. Voluntariness is not a defense against a sexual harassment charge; a person might submit to sexual advances due to fear of an unfavorable tangible job action such as termination.

The U.S. Supreme Court defined two types of sexual harassment. The first of these is "quid pro quo," where sexual involvement is explicitly or implicitly made a condition of employment or where submission or rejection of harassing conduct is used as a basis for employment decisions about the employee. The second type of harassment is "hostile work environment," where the conduct has the purpose or effect of interfering with the employee's work performance.

More recently, the courts have formulated standards for assessing employer liability in sexual harassment cases. A tangible employee action constitutes a significant change in employee status, such as hiring, firing, failing to promote, or reassigning with significantly different responsibilities, or a decision causing a significant change in benefits. If the alleged harasser is a supervisor and there is a tangible employee action, the case is considered quid pro quo and the employer is potentially vicariously liable. If there is no such action but the harassment is severe or pervasive, the employer may present an affirmative defense by showing that reasonable care was exercised to prevent sexually harassing behavior and to ensure that such behavior was corrected promptly and that the employee failed to take advantage of preventive or corrective opportunities or to otherwise avoid harm. There are variations of these rules for alleged harassment by coworkers or outsiders such as vendors associated with the employer.

Many employers have established firm policies to prevent and correct such harassment. At a minimum, these policies involve accessible and effective complaint processes, training for all employees, and detailed investigative procedures. The sensitivities involved in carrying out these policies and the possibility of sensational disclosures require the application of psychological knowledge. During the mid-1990s, Congress extended the so-called federal rape shield law to civil proceedings. Thus, a woman's current and past sexual behavior is admissible in court only if the defendant can show that its evidentiary value outweighs any prejudicial effect on the plaintiff. Thus, sexual harassment cases, which are civil cases, may

involve all of the applications of psychology now applied in rape cases. As would be expected, there are a number of legal gray areas in sexual harassment cases. Perhaps this is one of the reasons why the courts and legislatures have been reluctant to extend current laws beyond relationships between persons of the opposite sex. Even disgusting and gross behavior, as has occurred in same-sex horseplay at work, has not been considered legally as sexual harassment. Civil rights laws are not being considered to be general civility codes.

10. CONCLUSION

Although there are other pertinent connections between psychology and employment discrimination, it is clear that psychological knowledge and methods can have a major role in preventing and understanding such discrimination. In addition, knowledge in other psychological areas that are not generally considered to be applied can inform activities involved in possible employment discrimination. Applied psychology has become more highly developed in many respects due to the research and debates spurred by the civil rights movement and the emphasis on employment discrimination. The database of psychological knowledge has been greatly enhanced by these efforts in applied psychology.

See Also the Following Articles

Prejudice and Discrimination ■ Sexual Harassment

Further Reading

Age Discrimination in Employment Act of 1967, as amended, 29 U.S.C. 621.

American Educational Research Association, American Psychological Association, & National Council in Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Authors.

American with Disabilities Act, 42 U.S.C. 12101 (1990).

Civil Rights Act of 1964, § et seq., as amended, 42 U.S.C. § 2000, et seq.

Cleveland, J. N., Stockdale, M., & Murphy, K. R. (2000). *Men and women in organizations: Sex and gender issues at work.* Mahwah, NJ: Lawrence Erlbaum.

Federal Glass Ceiling Commission. (1995). *Good for business: Making full use of fact-finding report of the Federal Glass Ceiling Commission.* Washington, DC: U.S. Department of Labor.

Griggs v. Duke Power Company, 401 U.S.C. 424 (1971).

Lockwood, N. (2003). *The aging workforce: The reality of the impact of older workers and eldercare in the workplace.* Alexandria, VA: Society for Human Resource Management.

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection procedures.* Bowling Green, OH: Author.

Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. 1607 (1978).

Employment Interviewing

Deborah L. Whetzel

Work Skills First Inc., Glen Allen, Virginia, USA

Michael A. McDaniel

Virginia Commonwealth University, Richmond, Virginia, USA

GLOSSARY

criterion-related validity The correlation between a predictor (e.g., test, interview) and the criterion that it is designed to predict (e.g., job performance).

critical incidents Structured stories about good and poor job performance, including the situation in which the incident occurred, the behavior describing the performance, and the result of the incident.

incremental validity The validity of one predictor over and above that of another predictor.

interrater reliability An estimate of the reliability of a measure assessed by the correlation between ratings made by two raters independently.

meta-analysis A set of statistical techniques used to summarize research results across studies to draw generalizable conclusions.

past behavior interview question A question in which the interviewee is asked to describe his or her performance in a similar, but more specific, situation (e.g., "Describe a time when you had to give bad news to a client. What was the situation? What did you do? What was the result?").

situational interview question A question that poses a situation to the interviewee and asks how he or she would likely perform in that situation (e.g., "What would you do if . . . ?").

This article is organized into two sections: overview and psychometric characteristics of the interview. The overview section provides a definition of interviews, a description of various levels of structure, and methods for enhancing the structure of the interview. It also provides examples of questions using the job of a civil aviation security inspector. The psychometric section describes research on reliability, validity, correlations with other constructs, subgroup differences, and incremental validity beyond general mental ability.

1. OVERVIEW

1.1. Definition

An interview involves interactions (face-to-face, telephone, or written) between an interviewer and an interviewee. An interview typically is initiated to achieve one or more objectives (e.g., obtaining data in response to a survey, disciplining employees, providing feedback as part of performance appraisal, career counseling, selecting employees from a pool of applicants) and occurs as part of a sequence of activities (e.g., reviewing resumes, interviewing a subset of candidates, and administering cognitive ability tests for the purpose of selecting applicants; reviewing training curricula and interviewing employees for the purpose of assessing training needs). The interview's objectives and the sequence of activities in which the interview takes place are known as the context of the interview. In this article, interviews are

described in the context of selection, that is, interviews used to select employees from a pool of applicants. In this setting, interviews typically focus on past, present, or future behavior, beliefs, opinions, attitudes, and/or convictions of applicants. Information about an individual's previous experience, education, and vocational aspirations reported during the interview, as well as behavior observed during the interview, is considered the content of the interview. Interviewers are required to interpret information and draw inferences about traits, attributes, attitudes, and skills that the applicants possess. These inferences often take the form of ratings on dimensions of personal attributes or rankings of applicants.

1.2. Levels of Structure

There are two broad types of employment interviews: structured and unstructured. Structured interviews have a fixed format or a fixed set of questions, and the questions are based on a job analysis. Conversely, unstructured interviews do not use a standard procedure for the administration of questions or scoring. In fact, questions may vary from applicant to applicant.

In 1992, Huffcutt reviewed the literature on interview structure and described four progressive levels of structure. Level 1 was the typical unstructured interview characterized by no constraints on the questions and a global evaluation of responses. Level 2 imposed limited constraints by specifying the topics to be covered by the questions and some degree of structure on response evaluation. Level 3 required the prespecification of questions, although applicants were not asked precisely the same questions because different interview forms were used or interviewers were allowed to choose among alternative questions and to probe responses to the specified questions. An example of this kind of interview is the behavioral patterned description interview as described by Janz in 1982. Responses were evaluated using some degree of structured response evaluation. Level 4 involved asking applicants precisely the same questions with no deviation or follow-up probes, and responses were scored according to benchmark answers. In 1980, an example was provided by Latham and colleagues, who discussed situational interviews.

Situational and past behavior are two primary types of structured interviews. Situational questions provide a situation to the interviewee and ask how he or she would likely perform in that situation (e.g., "What would you do if . . . ?"). The situations resemble those in situational judgment tests (SJTs). Past behavior questions provide a broad situation for the interviewee.

The interviewee is asked to describe his or her performance in a similar, but more specific, situation (e.g., "Describe a time when you had to give bad news to a client. What was the situation? What did you do? What was the result?"). Both of these kinds of questions derive from critical incidents provided by subject matter experts (SMEs). Critical incidents are structured stories about good and poor job performance. The three-part structure of critical incidents involves having the SMEs respond to three questions: "What was the situation/task?," "What did you do?," and "What was the result?"

Interviews can be very appealing to organizations because questions can be created to tap cognitive and noncognitive skills as well as abilities. Interviews can focus on a range of applicant qualities such as communication skills, interpersonal skills, decision making, and substantive knowledge. Examples of situational and past behavior interviews are provided in Tables I and II, respectively.

1.3. Methods for Enhancing the Structure of Interviews

In 1997, Campion and colleagues reviewed the use of structure in selection interviews and identified 15 components of structure that may enhance either the content of the interview or the evaluation process in the interview:

1. Base questions on a job analysis.
2. Ask the exact same questions of each candidate.
3. Limit prompting, follow-up questioning, and elaboration on questions.
4. Use better types of questions (e.g., situational, past behavior, background, job knowledge).
5. Use a longer interview or a larger number of questions.
6. Control ancillary information.
7. Do not allow questions from the candidate until after the interview.
8. Rate each answer or use multiple scales.
9. Use detailed anchored rating scales.
10. Take detailed notes.
11. Use multiple interviewers.
12. Use the same interviewer(s) across all candidates.
13. Do not discuss candidates or answers between interviews.
14. Provide extensive interviewing training.
15. Use statistical prediction rather than clinical prediction.

Critical Incident

During an air carrier assessment, an inspector noticed that the screeners at checkpoints appeared to be unsure about the proper use of metal detector equipment and procedures for screening passengers. The inspector interviewed all of the screeners on-site and discovered that the screeners on duty at the metal detectors had not received required training. The inspector immediately informed all air carriers represented at the checkpoint and requested that carriers implement additional screening procedures. The screeners who had not been trained were removed immediately.

Situational Question

You are conducting an air carrier assessment and you notice that several of the screeners at the metal detectors appear to be unsure of how to operate the equipment. What would you do?

Response Anchors to Situational Question

Highly ineffective	*Average*	*Highly effective*
• Have the inspector on the next shift train the screeners • Do nothing, assuming that the screeners will get better using the machinery with practice (e.g., learning by doing)	• Tell another inspector about the problem • Watch for multiple instances in which screeners appear to be unsure of how to operate the equipment to make sure that you are correct in your assessment	• Inform the air carriers being represented at the checkpoint • Have the untrained screeners removed and/or trained immediately

1	2	3	4	5	6	7	8	9

These components of structure are based on research literature and continue to be replicated. The more of these components that are incorporated into the interview, the more structured the interview becomes.

2. PSYCHOMETRIC CHARACTERISTICS OF THE INTERVIEW

2.1. Reliability

An interview is reliable to the extent that it is free of measurement error. Typically, reliabilities of interviews are assessed by interrater reliability, that is, the correlation of the interview ratings made by two independent raters. There are at least two meta-analytic reviews that investigated the reliability of interviews. In 1995, Conway and colleagues found the average reliability of structured interviews to be .67. In 1994, McDaniel and colleagues found mean reliabilities of .84 for job-related structured interviews (across 167 coefficients) and .68 for job-related unstructured interviews (across 20 coefficients).

2.2. Validity

The validity of the interview is typically assessed by correlating an interview score with a measure of job performance. This is one application of criterion-related validity. McDaniel and colleagues broke down interviews by content and found that situational interviews were superior to job-related and psychological interviews with respect to job performance (criterion-related validity coefficients of .50, .39, and .29, respectively). Structured interviews, regardless of content, were more valid (.44) than unstructured interviews (.33) for predicting job performance criteria. Individual interviews (.43) were more valid than board interviews (.32).

In 1995, Pulakos and Schmitt found that experience-based interview questions yield higher validities (.32) than situational questions (–.02). This result was replicated by Huffcutt, Weekley, and colleagues in 2001. Their behavioral description interviews (BDI) (i.e., experience-based interviews) were a reasonably strong predictor of leadership performance in training for Canadian military officers, whereas their total situational interview (SI) scores were not ($r = .49$ vs .22 after correcting for range restriction). This result was replicated again by Huffcutt and colleagues using a

TABLE II
Past Behavior Interview Example

Critical Incident

A senior inspector wanted on-the-job training (OJT) for new inspectors in the office on the procedure used by the legal counsel to close a case or hold informal hearings. The senior inspector contacted the legal counsel's office and discussed the matter with several lawyers who handled security cases. The lawyers agreed to provide OJT for the new inspectors. Consequently, new inspectors now routinely attend informal conferences with legal counsel and receive firsthand instruction on how informal hearings are conducted and how to close a case.

Past Behavior Question

Tell me about a time when you needed to have people trained as soon as possible and you did not have the expertise to provide the training yourself. What did you do?

Response Anchors to Past Behavior Question

Highly ineffective	*Average*	*Highly effective*
• Have new inspectors "learn it on their own" on the job • Perform work on own that could be delegated if inspectors were trained • Give more assignments to top performers rather than train new Inspectors	• Read a book and learn the subject area yourself • Identify and suggest training needs for new inspectors in a variety of areas of expertise • Coach new inspectors as they learn new things	• Identify individuals who are knowledgeable in the subject area and ask them to teach the new inspectors • Provide developmental opportunities to new inspectors who wish to learn a new function • Create opportunities for inspectors to try things they have not done before

1	2	3	4	5	6	7	8	9

sample of district managers. The BDI total scores correlated with the performance evaluation ($r = .31$), whereas the SI total scores did not ($r = .02$).

2.3. Correlations with Other Constructs

In 2001, Huffcutt, Conway, and colleagues developed a taxonomy of seven types of constructs likely to be measured by interviews. Analysis of 338 ratings from 47 actual interview studies indicated that basic personality and applied social skills were the most frequently rated constructs, followed by mental capability and job knowledge and skills. These authors also found that high- and low-structure interviews tend to focus on different constructs. Furthermore, they suggested that the reason why structured interviews tend to have higher validity is because they focus more on constructs that have stronger relationships with job performance. High-structure interviews focus more on applied mental skills, direct job knowledge, applied social skills (e.g., communication, interpersonal skills, leadership) and organizational fit, whereas low-structure interviews focus more on general mental ability, background credentials (e.g., education, training, experience), some aspects of

personality (e.g., agreeableness, emotional stability), and physical attributes. For Huffcutt and colleagues' analysis across all studies, the validity coefficients ranged from .24 to .58. As expected, the mean validities for constructs rated in more structured interviews were higher (.39) overall than the mean validities for constructs rated in less structured interviews (.24).

Two meta-analytic studies documenting the relationship between interview ratings and cognitive ability scores are described later.

2.4. Subgroup Differences

Most employment-related assessments show race differences in mean scores. On average, White job applicants obtain higher mean scores than Hispanic job applicants, who in turn have higher mean scores than Black job applicants. These mean score differences are typically large enough to sharply curtail the number of Blacks selected for employment. Mean race differences in employment interview scores are often smaller than other measures because the interview is often the final selection procedure used in a multiple-procedure process. For example, if a written job knowledge test is used, Blacks and Whites will likely show large mean

score differences. However, if only those who score high on the job knowledge test are interviewed, the Blacks and Whites who are interviewed will be more similar in job-related skills than will the pool of all applicants who took the job knowledge test. This effect, known as "range restriction," will result in an interview showing small or no race differences. Pulakos and Schmitt found very small mean differences between subgroups on interview scores. However, they used job incumbents in both of their samples, suggesting that selection to enter the job may have absorbed most of the adverse impact. Pulakos and Schmitt generally concluded that interviews yielded fewer subgroup differences and widely recommended their use. However, in 2002, Roth and colleagues found that two forms of a behavioral interview were associated with standardized ethnic group differences of .36 and .56 when corrected for range restriction. These differences were three to five times as large as the previously estimated value of .10 obtained by Huffcutt and Roth in 1998. Roth and colleagues attributed these differences largely to the fact that applicants were previously screened on measures with high subgroup differences (e.g., cognitive ability), similar to the Pulakos and Schmitt study. They recommended that when examining the psychometric properties of a predictor, analyses should focus on the applicant level, especially applicants who have not been prescreened on another predictor.

2.5. Incremental Validity beyond General Mental Ability

In 2000, Cortina and colleagues found that interview scores contribute to the prediction of job performance over and above cognitive ability and conscientiousness to the extent that they are structured, with scores from highly structured interviews contributing substantially to prediction. Specifically, they found that the incremental validity for Level 1 (lowest level of structure) interview scores ranged from 0.9 to 2.2% of the variance in job performance. Level 2 (moderate level of structure) interview scores explained an additional 1.8 to 6.2% of the variance in job performance. Level 3–4 (highest level of structure) interview scores explained an additional 12.3 to 22.2% of the variance in job performance over and above cognitive ability and conscientiousness. Pulakos and Schmitt also found that interviews had incremental validity beyond mental ability.

In 1996, Huffcutt and colleagues explored the extent to which employment interview evaluations reflect cognitive ability. A meta-analysis of 49 studies found a corrected mean correlation of .40, suggesting that on average 16% of the variance in interview constructs represents cognitive ability. In fact, their results suggest that interview ratings that correlate higher with cognitive ability tend to be better predictors of job performance.

3. CONCLUSION

Employment interviews are frequently used selection tests. Employment interviews vary in their degree of structure. Structured interviews permit a more systematized collection of job applicant information than less structured interviews, resulting in the former providing better predictions of future behavior. The reliability of interviews is typically assessed by correlating the interview ratings of two independent interviewers. Interview structure tends to enhance reliability. Interviews can be used to measure a variety of job-related personal characteristics, including personality, applied social skills, job knowledge, and cognitive ability. Like most selection procedures, employment interviews tend to show mean differences in scores as a function of race. There is some evidence that interviews provide incremental prediction beyond that of general mental ability. In general, it is recommended that employment interviews be structured and based on knowledge of the job.

See Also the Following Articles

Competence at Work ■ Cooperation at Work ■ Job Analysis, Design, and Evaluation ■ Work Motivation ■ Work Role, Values Sought in the

Further Reading

Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology*, *50*, 655–701.

Conway, J. M., Jako, R. A., & Goodman, D. F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, *80*, 565–579.

Cortina, J. M., Goldstein, N. B., Payne, S., Davison, H. K., & Gilliland, S. W. (2000). The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology*, *53*, 325–351.

Huffcutt, A. I. (1992). *An empirical investigation of the relationship between multidimensional degree of structure and the*

validity of the employment interview. Unpublished doctoral dissertation, Texas A&M University.

Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology*, 86, 897–913.

Huffcutt, A. I., & Roth, P. L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology*, 83, 179–189.

Huffcutt, A. I., Roth, P. L., & McDaniel, M. A. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity. *Journal of Applied Psychology*, 81, 459–473.

Huffcutt, A. I., Weekley, J. A., Weisner, W. H., DeGroot, T. G., & Jones, C. (2001). Comparison of situational and behavior description interview questions for higher-level positions. *Personnel Psychology*, 54, 619–644.

Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology*, 67, 577–580.

Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology*, 65, 422–427.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599–616.

Pulakos, E. D., & Schmitt, N. (1995). Experience-based and situational interview questions: Studies of validity. *Personnel Psychology*, 48, 289–308.

Roth, P. L., Van Iddekinge, C. H., Huffcutt, A. I., Eidson, C. E., & Bobko, P. (2002). Corrections for range restriction in structured interview ethnic group differences: The values may be larger than researchers thought. *Journal of Applied Psychology*, 87, 369–376.

End of Life Issues

Rebecca S. Allen, Laura L. Phillips, and Kristen Payne

University of Alabama, Tuscaloosa, Alabama, USA

GLOSSARY

advance care planning (*ACP*) The process of preparing for medical status change events at the end of life.

advance directive A legal document that explains a competent individual's wish for future medical care (as in a living will) or that designates a surrogate decision maker who can make medical treatment decisions for the individual when he or she is unable to communicate these treatment preferences independently (as in a durable power of attorney for health care, also known as a health care proxy).

anticipatory grief A process of grieving for losses encountered during the course of illness and prior to death.

hospice care Palliative and supportive services that provide physical, psychological, social, and spiritual care for dying persons and their families; patients must have a prognosis of 6 months or less to live, receive no life-prolonging medical treatments, and have an identified primary caregiver.

palliative care Extends the principles of hospice care to a broader population who can benefit from receipt of aggressive comfort care earlier in the disease process than the point at which life expectancy is 6 months or less; provision of palliative care is not exclusive of receipt of life-prolonging medical treatments or independent living situations.

treatment delivery An aspect of treatment implementation that pertains to the degree to which project staff adhered to treatment guidelines and documented their treatment package (i.e., development of interventionist training manuals).

treatment enactment An aspect of treatment implementation that pertains to the usefulness of the treatment to the individual and/or family as indicated by their spontaneous use of treatment skills in everyday life.

treatment receipt An aspect of treatment implementation that pertains to the individual and/or family's understanding of treatment components and acquisition of skills.

Everyone must deal with the deaths of those they love as well as his or her own mortality. Yet one rarely considers the certainty of the end of life until confronted with the immediate needs of a loved one who is dying. This article reviews the differentiation of hospice and palliative care, predictors of mortality, advance care planning within a family context, caregiving stressors as well as positive aspects of caregiving, and assessment and treatment issues surrounding the end of life.

1. INTRODUCTION

Issues surrounding the end of life are receiving increased attention due to the "graying of America." In 2000, the average life expectancy for all individuals at birth in the United States, regardless of gender or race, was 77 years. The National Center for Health Statistics reported that the leading causes of death in 2001 were heart disease, cancer, stroke, and respiratory disease. Alzheimer's disease was listed as the eighth leading cause of death, although this is likely an underestimate given that Alzheimer's is rarely listed as the direct cause of death on death certificates. Fully 80% of individuals who die do so in an institutional setting such as a hospital or nursing home. However, individuals report a desire to die at home.

One reason for the increased interest in end of life issues has to do with the U.S. Supreme Court's 1990 decision in *Cruzan v. Director, Missouri Department of Health,* requiring "clear and convincing evidence" of an individual's wishes regarding medical treatments that may be used to prolong life (e.g., artificial nutrition and hydration). Many cultural and psychosocial issues influence individual preferences for end of life medical treatments, receipt of hospice or palliative care, nursing home placement, and the experience of caregiving. These include a need for education regarding the benefits and drawbacks of life-prolonging (or life-sustaining) treatments in comparison with receipt of hospice and palliative care, socioeconomic factors, religiousness and spirituality, prior experience with end of life issues, and style of family decision making.

2. DEFINITIONS AND DIFFERENTIATION OF HOSPICE AND PALLIATIVE CARE

Hospice care includes palliative and supportive services that provide physical, psychological, social, and spiritual care for dying persons and their families. The National Center for Health Statistics reported that hospice patients in the United States are still predominantly cancer patients (e.g., 58% in 2000). However, guidelines to determine hospice eligibility have been developed for severe chronic illnesses such as heart disease, chronic obstructive pulmonary disease, cerebrovascular disease, and dementia. There are dramatic racial/cultural differences in the receipt of hospice care. Hospice patients are most often elderly White individuals who live with a primary caregiver in a private or semiprivate residence and are covered by Medicare or private insurance. Only 5% of hospice services are paid for by Medicaid. Although the Medicare hospice benefit provides for care during the final 6 months of life, the average length of service in hospice programs nationally is only 46.9 days (median = 15.6 days).

A primary restriction to the provision of hospice care is that patients must have a life expectancy of 6 months or less. In addition, no life-prolonging treatments (e.g., artificial nutrition and hydration, chemotherapy, ventilation) can be ongoing during the provision of hospice care. This may be a primary barrier to receipt of hospice services by individuals who identify themselves as members of racial/ethnic minority groups. For example, African Americans are more likely to initiate life-prolonging medical treatments than are Whites. Another barrier to receipt of hospice services is the absence of an informal or formal in-home caregiver.

In contrast, palliative care extends the principles of hospice care to a broader population who could benefit from receiving aggressive comfort care earlier in the disease process. Palliative care is not restricted to individuals with a life expectancy of 6 months or less. Furthermore, provision of palliative care is not exclusive of receipt of life-prolonging medical treatments or independent living situations. Therefore, palliative care models are ideal for more medically frail elders who may benefit from aggressive comfort care well before they qualify for current hospice guidelines. Ideally, palliative care would segue into hospice care as the individual's illness progresses. However, recognizing the signs of disease progression that can signal a need for palliative care can be difficult. Recent research has attempted to clarify some of these signs and symptoms.

3. PREDICTORS OF MORTALITY

Events that can signal medical status change include a nonchronic condition that calls for medical follow-up by a physician (e.g., weight loss, difficulty in breathing, heart trouble, chest pain, and trouble in swallowing) and a disease process that calls for a medical decision involving the individual's family. In 2003, Covinsky and colleagues examined data from the final 2 years of life of frail older adults in the community and found that declines in functional status were evident at least 1 year prior to death. The degree of functional impairment was greater for those with cognitive impairment.

Medically frail individuals are at risk for various types of cognitive impairment, including dementia, confusion associated with opioid medication, delirium, and (in the case of cancer) brain metastases. In 2003, Schonwetter and colleagues found that advanced age and impaired nutritional and functional status were associated with shortened survival among patients admitted to a community hospice with a diagnosis of dementia. Three-fourths of patients meeting local Medicare fiscal intermediary guidelines for hospice eligibility died within the required 6-month window.

In addition, continued weight loss has been associated with an increase in the likelihood of dying. Weight loss is a common symptom in many types of cancer, although fluid buildup is not uncommon. Older age and feeding dependence are also associated with increased risk of mortality. However, the most prevalent symptom of impending death during the final 48 hours appears to be respiratory symptoms.

Perhaps the most common symptom of medical status change, and certainly the most troubling for the patient and the family caregiver, is pain. Individuals with more severe cognitive impairments might not be reliable in, or even capable of, self-reporting pain or discomfort. Individual pain reports might not coincide with proxy pain reports. Pain is common among nursing home residents regardless of proximity to death. In 2002, Happ and colleagues noted that most nursing home residents died without family members present and with little documented evidence of pain or symptom management. According to Engle and colleague's 1998 research, African American nursing home residents may be at particular risk for inadequate pain relief.

The National Hospice and Palliative Care Organization (NHPCO) has recommended continuous assessment of an individual's needs and treatment preferences in the context of family values. This process may be considered continual treatment planning for the end of life, and some have argued that this should occur within a family context.

4. ADVANCE CARE PLANNING WITHIN A FAMILY CONTEXT

Advance care planning (ACP) is the process of preparing for medical status change events at the end of life. ACP typically involves discussion and assessment of the understanding that the person with life-limiting illness has regarding his or her medical history, current condition, values, preferences, and resources. ACP tends to emphasize decisions regarding the withholding of life-prolonging medical treatments rather than treatments that are desirable. It is important to note that ACP is not merely the completion of an advance directive, although completion of an advance directive may be a part of ACP.

Older adults tend to believe that others will take care of their future health care. This belief may, in turn, preclude completion of advance directives or interfere with ACP. Individuals who possess advance directives tend to be older, White, more educated, classified as middle class, and in poor health.

ACP is typically conducted in a family context because family members are often involved in end of life care. Family members tend to be present throughout the individual's illness and play an important role in end of life decisions. Sometimes, family members advocate the individual's end of life treatment preferences when he or she cannot voice these wishes independently. However, this substituted judgment is not always an accurate representation of the individual's wishes. This discrepancy may be influenced by the lack of prior discussions of end of life preferences with caregivers or family members; however, a growing body of research suggests that even with such discussions, surrogate decision makers do not always respond with the same treatment preferences as indicated by their loved ones. Many factors may contribute to this discrepancy, including differences in how healthy and unhealthy individuals view illness and quality of life. Although research is still inconclusive as to whether ACP and completion of advance directives aid in maintaining patient control over the dying process, there seems to be some suggestion that at least ACP may ease the minds of patients, caregivers, and family members concerning end of life issues.

5. FAMILIAL CAREGIVER STRAIN AND GROWTH POTENTIAL

On the topic of caregiving, research to date has focused on caregiving for chronic illnesses, most predominantly Alzheimer's disease and other forms of dementia. However, there are differences in the impact of caregiving across various diseases and racial groups. When reviewing research on caregiving at the end of life, important themes emerge, including the relationship between caregiver and patient; the search for meaning/purpose, spirituality, or religiousness; and the adjustment to bereavement.

In general, caregivers experience more negative health outcomes than do noncaregivers. Specifically, caregivers are at increased risk for developing negative mental and health outcomes such as depression, perceptions of burden, lowered immune system function, and mortality. Female caregivers experience less participation in social and recreational activities, and more movement out of the workforce, than do noncaregivers. Research has suggested that Alzheimer's disease produces similar negative effects on caregivers' well-being and mental health when severity of illness is considered.

Other important aspects to consider when evaluating the long-term effects of caregiving are the quality and type of the relationship between caregivers and the individuals for whom they are providing care. In 1998, Lawrence and colleagues found that better relationship quality was directly related to lower levels of depression and a decreased sense of role captivity for caregivers. Familial relationship has been found to influence the average number of hours per week spent in the caregiving role and the extent of burden experienced. Specifically, spouses tend to spend more time caregiving than do adult children or other family members. In addition, spousal caregivers report more fatigue and more sleep difficulties than do nonspousal caregivers.

Regarding racial differences, African Americans are more likely to be cared for by members of their extended families than are Whites, whereas adult children are a prominent source of care for both African Americans and Whites. African American caregivers also experience higher levels of overall coping mastery than do White caregivers. In contrast, Aranda noted in 2001 that Hispanic/Latino caregivers generally, but not always, report higher levels of depressive symptoms than do other ethnic/racial groups.

Beyond the impact of type of relationship or the length of care provided, the course of the illness also affects the relationship and caregivers' response to caregiving. Although a good deal of research has focused on caregiving for chronic illness, very little has focused on caregiving at the end of life. Most of what is known in this area comes from clinical reports of the dying process or observations of experienced clinicians who work with families through hospice and palliative care programs. These qualitative accounts portray the extremely difficult circumstances experienced at the end of life by caregivers, who are simultaneously caring for a loved one while also dealing with the anticipatory grief related to their loved one's impending death as well as the frustrations of dealing with an often unresponsive health care system.

Notably, few studies address potential positive aspects of palliative caregiving for persons at the end of life. Caregivers may face needs that could be identified as existential or spiritual, including finding purpose and meaning, forgiving and receiving forgiveness, maintaining hope, saying good-bye, and coming to terms with expectations for what may occur after death. In 2003, Haley and colleagues found that hospice caregivers who have high levels of social activity and social support, and who are able to find meaning and benefit in caregiving, reported lower levels of depression and higher levels of life satisfaction. This recent finding suggests that activities in which caregivers find meaning and benefit in caregiving are beneficial.

The impact of religiousness and spirituality on coping has been a particularly neglected area of research. Religiousness and spirituality may reduce the sense of loss of control and helplessness that can accompany end of life caregiving. Religious and spiritual beliefs provide a cognitive framework that reduces stress and increases purpose and meaning in the face of illness. Spiritual and religious activities may reduce the sense of isolation and increase caregivers' sense of forgiveness or control over the illness. However, religiousness and spirituality might not always have a positive influence. For example, if a person who is dying feels remorse over past failings and believes in a punishing God, he or she may experience great anxiety.

When death comes after a sustained period of chronic illness and family caregiving, family members often react with a mixture of relief, grief, and guilt. Relief at the end of the caregiving process and the end of the loved one's suffering is accompanied by a deep sense of loss. Family members experience increased stress following a decision to terminate end of life care, but in cases where there is no verbal or written advance directive, the stress levels tend to be higher and last longer. Family members may experience anticipatory grief during their loved one's dying process. This anticipatory grief process may be beneficial to adaptive coping after bereavement, but this has not been supported consistently by research.

Although most people go through the grieving process effectively and eventually resume their normal functioning, a considerable portion (perhaps 10–25%) experience some form of pathological grief. A number of risk factors for complications with bereavement have been found, including the quality of the relationship between the bereaved individual and the deceased loved one (worse outcome in conflictual relationship), gender (worse outcome in male caregiver), and the

nature of the death (worse outcome in sudden, unexpected, or socially unacceptable death). The impact of caregiving does not end at death, nor is this impact the same for every caregiver. Substantial racial/ethnic differences in subjective reactions to death have been found. Specifically, African American dementia caregivers are more likely than White dementia caregivers to initiate life-prolonging medical treatments, are less likely to have their loved one die in a nursing home, and report less acceptance of their loved one's death and greater perceived loss. Consequently, more research is needed to understand the impact of caregiving, and an individualized approach to working with these families in clinical settings is needed.

6. THERAPEUTIC ASSESSMENT

Careful assessment is a precursor to effective treatment, and effective treatment leads to enhanced quality of life for individuals with life-limiting illnesses. The importance of fine-tuned information on patient prognosis becomes more apparent as treatment options increase. Communication of such information to individuals and their families allows time to explore family values and goals for future medical treatment. Such medical treatment can be viewed as life sustaining or life prolonging, depending on the perspective of each individual and family.

Assessment of patients' preferences regarding future medical treatment can be conducted formally through the use of advance directives or informally through discussions with family members. One assessment tool that may be helpful in promoting informal discussions and in clarifying wishes for future medical care is the Life Support Preferences/Predictions Questionnaire (LSPQ). This instrument was designed to identify treatment preferences in specific illness scenarios. Each scenario differs from the others in severity, level of pain, prognosis, and/or impairment. Following presentation of each scenario, respondents are asked to indicate which treatment options they would desire given that situation. Data from such assessments can provide guidelines for appropriate treatment and disease management.

Clinicians also may wish to conduct needs assessments with individuals and family members throughout the course of the individuals' chronic life-limiting illnesses. Continuous and comprehensive assessment of individual and family perception of need can provide a medium for matching specific individual and family characteristics with appropriate therapeutic interventions. Data from such assessments can provide education and support for individuals living with illnesses at the end of their lives. Such information can also provide education and support for family members regarding their own risk for disease and available preventive measures.

7. THERAPEUTIC INTERVENTIONS AND TREATMENTS

Many families and individuals with life-limiting illnesses have difficulty in accepting the inevitable consequences of the diseases and may resist engaging in interventions such as advance care plans, psychotherapy, support groups, and respite services or seeking information about the benefits of palliative care to alleviate suffering. Clinicians must recognize the potential of such resistance and approach individuals and families with the utmost sensitivity and empathy. However, once individuals and families are engaged in discussions of therapeutic interventions at the end of life, the opportunity for meaningful interaction and the potential for individual growth and growth in relationship become more apparent. Some interventions that target various stages of life-limiting illnesses and end of life treatment issues are reviewed briefly in the following subsections.

7.1. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments

The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT) project, conducted in 1995 by the SUPPORT principal investigators, implemented and evaluated a physician-assisted intervention for dying patients in hospitals. The intervention was aimed at addressing inadequate communication regarding prognoses and patient care preferences. Specifically, the intervention was aimed at (a) providing accurate and timely prognostic information to physicians, (b) documenting the views of patients and families surrounding the patients' prognoses and treatment options, and (c) providing skilled nurses to carry out discussions and communicate relevant information.

To assess whether the intervention had a positive effect, SUPPORT researchers focused on (a) when and if "do not resuscitate" (DNR) orders were written, (b) patient–physician agreement regarding preferences to withhold resuscitation, (c) the number of days patients spent in intensive care units, (d) the frequency and

severity of the pain experienced by patients, and (e) the amount of hospital resources used during patients' hospital stays. The research investigators found no differences on any of the five outcomes between those who received the intervention and those who did not.

The specific reasons why the intervention did not work are unclear, but there are some suggested possibilities. The SUPPORT investigators recognized that one limitation in their ability to detect a positive impact of this treatment might be that individuals, families, and physicians were generally satisfied with the care received at that critical time. According to the researchers, another possible explanation is that the hospitals in which the SUPPORT study was conducted might have provided better care than is typically available to the general public.

Although these are plausible explanations, another possibility is variability in the implementation of the intervention itself in terms of treatment delivery by the investigators, treatment receipt by the participants (i.e., understanding), and treatment enactment (i.e., the degree to which the information provided by the SUPPORT team was used by the families in the study). SUPPORT was the first intervention of its kind to submit to the rigors of scientific testing and represents seminal research in this area. The investigators chose to maximize generalizability by conducting the research in as natural a way as possible in their chosen hospitals. However, in so doing, the experiences of individuals and families in the treatment group varied dramatically.

Delivery of the treatment was allowed to vary based on the SUPPORT nurse's discretion and ranged from informational support to intense emotional support. In addition, the nurse did not always meet with the families or the patients, nor did the nurse discuss the same topics with each participant. This variation in treatment can be beneficial to the individuals involved, but it makes it difficult to determine whether the treatment as a whole was effective or ineffective due to the variety of interventions delivered.

The only measure of receipt of treatment was the physicians' indications of receipt of the reports regarding DNR or the placement of the reports in the medical charts. Although the placement of these reports in the charts could be conceptualized as treatment enactment, there is little information on whether these reports were ever actually used by those in constant contact with the patients or families. Efforts to standardize the treatment provided by the interventionists, measures of treatment receipt for both physicians and patients/families, and a measure addressing the use (enactment) of the

treatment are needed before such interventions can be deemed beneficial or ineffective.

7.2. Advance Directives, Values Assessment, and Communication Enhancement

In 2001, Ditto and colleagues implemented an intervention that consisted of patients completing an advance directive in the presence of family members. In some groups, discussions were also encouraged to facilitate communication between patients and family members regarding medical treatment preferences. One group had advance directives with discussions, one group had advance directives without discussions, one group had discussions without advance directives, and a control group had neither advance directives nor discussions. The results of this study showed that none of the interventions resulted in an increase in patient–family member agreement when family members were asked to predict patient preferences. Although the study revealed no increase in agreement, it did reveal an increase in comfort and understanding as reported by patients and family members regarding end of life decisions. As with the SUPPORT study, data regarding treatment implementation were not reported systematically, although there seemed to be variability in the delivery of treatment to the various groups.

7.3. Meaning-Centered Group Interventions

William Breitbart and colleagues are currently testing an intervention designed to enhance a sense of meaning and purpose and to improve quality of life among advanced cancer patients. The treatment is delivered in a group format (eight sessions) and targets the creation of a supportive environment to facilitate understanding of the possible sources of meaning before and during cancer illness. Basic concepts include the achievement of transcendence through creativity (e.g., work, deeds, causes), experience (e.g., nature, art, relationships), and attitudes toward suffering. Group sessions include information about sources of meaning, a storytelling life project, and hopes for the future. Anecdotal reports from Breitbart and colleagues indicate that the intervention may be effective in decreasing hopelessness and desire for hastened death while increasing meaning and spiritual well-being. No treatment implementation data are yet available.

7.4. Individual Problem-Solving Skills Training

Another type of intervention used at the end of life is problem-solving skills training. Susan McMillan's "Caregiver Intervention to Improve Hospice Outcomes" is one example of a problem solving-based intervention. This intervention takes place over a period of 9 days, during which time three intervention visits occur. In these visits, caregivers receive the "COPE" intervention involving the following: Creativity (viewing problems from different perspectives to find new strategies), Optimism (having a positive but realistic attitude toward the problem-solving process), Planning (setting goals and thinking out steps to reach those goals), and Expert information (learning what caregivers need to know about the patient's condition, when to seek professional help, and what can be done by caregivers to help the patient).

The three problem areas on which McMillan focused regarding the hospice patient's condition were pain, constipation, and dyspnea. For each of these problems, a specific module was designed for caregivers to learn assessment and problem-solving techniques they could implement to help the patient. In addition, caregivers were asked to keep a patient symptom diary in which they recorded patient symptoms to monitor and share this information with the hospice staff as needed.

Treatment implementation was measured in this study. Delivery of the treatment was monitored by case conferences to reduce the possibility of drift from the protocol or cross-contamination of the treatment conditions. In addition, visits in the intervention group were tape-recorded, and 10% of the tapes were reviewed monthly to ensure adherence to protocol. At monthly case conferences, the possibility of an undesired change in hospice staff behavior was assessed by a question regarding any noticed changes by caregivers with respect to the hospice staff's behavior toward the patient.

Treatment receipt, or caregivers' understanding of the intervention, was verified by the interventionist's inquiry, "What would you do if …?" Clarification regarding specific skills was provided to caregivers if needed to facilitate adequate treatment receipt. At the end of the intervention, receipt was assessed by asking caregivers to rate their ability to manage symptoms successfully and comparing their responses prior to the intervention with their responses following the intervention.

Treatment enactment can be examined through the caregivers' diary to determine the extent to which caregivers complied with the daily assessment and recording of patient symptoms. If caregivers found the information and monitoring system provided by the project to be helpful, they would be expected to use the skills in providing care to the patient.

7.5. Ongoing Studies by the Authors

Our team has developed two interventions that combine problem-solving skills training with social support and meaning-centered therapeutic activity. One intervention is the care integration team (CIT), which focuses on caregivers of terminally ill patients enrolled in hospice care. The intervention consists of four weekly home visits and has a randomly matched minimal support treatment control group. The intervention for caregivers focuses on use of problem-solving skills in four topic areas: (a) symptom management, (b) communication with hospice staff, (c) caregiver self-care, and (d) anticipatory grief. Treatment implementation measures are included in this ongoing study.

Unlike the CIT intervention, the legacy intervention focuses on both the individual with life-limiting illness and the caregiver. The legacy intervention is an effort to help patients and caregivers to find meaning through intergenerational activities. This is achieved through the joint development of the patient's "personal legacy" for future generations. This legacy can take multiple forms: a family photo album, a family cookbook or scrapbook, audiotaped or videotaped stories of the older adult's life, personal letters to younger family members, and the like. The intervention consists of three weekly 60- to 90-minute sessions in which the caregiver and patient learn about what a legacy is, select a legacy, obtain materials to create the project, and review their experience with members of the research team. A control group receives weekly 5-minute telephone calls and answers questions about spontaneous intergenerational activities to promote family ties. Encouraging families to develop lasting memories with their loved one at any stage of the disease process can enhance meaning in relationships and provide a focus to heal prior conflicts.

These interventions were designed with the three components of treatment implementation in mind. Both the CIT and legacy projects use audiotapes of intervention sessions (10%) in evaluations by an independent professional clinician not involved with the research project. The independent evaluator monitors the tapes and transcripts for interventionists' adherence to research protocol in treatment delivery. Both projects have caregiver and interventionist training manuals that are provided as guides to treatment components.

The CIT intervention measures treatment receipt by knowledge assessments at the first week of training and immediately following the intervention. If caregivers do not achieve an 80% score on these assessments, further training is provided to ensure receipt of the skills. Both the CIT and legacy interventions also use contact logs to document the content and length of all sessions so as to measure receipt of treatment, including who initiates the contact and the purpose of the contact.

The CIT project assesses treatment enactment by interventionists' review of caregivers' personal tracking forms. The interventionists then rate caregivers' compliance with all aspects of treatment on a Likert-type scale evaluation. For treatment enactment, interventionists working on the legacy project note the extent to which a legacy was completed by the family, again using a Likert-type scale evaluation.

8. CONCLUSION

This article has reviewed issues associated with personal and familial stressors and potential for growth when facing end of life issues during adulthood. Because of increased numbers of individuals living to very old age today, research and clinical attention are being directed at end of life care. Of necessity, end of life issues such as ACP, choice of medical treatment, assessment of treatment needs, and delivery of therapeutic treatments involve both individuals with life-limiting illnesses and their family members. Knowing when to initiate aggressive curative or aggressive palliative treatments is difficult given variable predictors of mortality across disease categories. This article has attempted to provide a broad overview of current clinical research efforts surrounding end of life issues. Without question, these topics will continue to be a focus for research, service, and policy efforts during the coming decades.

Acknowledgments

See Also the Following Articles

Further Reading

Allen, R. S., DeLaine, S., Chaplin, W. F., Marson, D. L., Bourgeois, M., Dijkstra, K., & Burgio, L. D. (2003). Advance care planning in nursing homes: Correlates of capacity and possession of advance directives. *The Gerontologist*, *43*, 309–317.

Allen, R. S., Haley, W. E., Small, B. J., & McMillan, S. C. (2002). Pain reports by older hospice cancer patients and family caregivers: The role of cognitive functioning. *The Gerontologist*, *42*, 507–514.

Allen, R. S., & Shuster, J. L. (2002). The role of proxies in treatment decisions: Evaluating functional capacity to consent to end-of-life treatments within a family context. *Behavioral Sciences and the Law*, *20*, 235–252.

Aranda, M. P. (2001). Racial and ethnic factors in dementia care-giving research in the U.S. *Aging and Mental Health*, *5*(Suppl. 1), S116–S123.

Covinsky, K. E., Eng, C., Lui, L. Y., Sands, L. P., & Yaffe, K. (2003). The last 2 years of life: Functional trajectories of frail older people. *Journal of the American Geriatrics Society*, *51*, 492–498.

Ditto, P. H., Danks, J. H., Smucker, W. D., Bookwala, J., Coppola, K. M., & Dresser, R. (2001). Advance directives as acts of communication: A randomized controlled trial. *Archives of Internal Medicine*, *161*, 421–430.

Engle, V. F., Fox-Hill, E., & Graney, M. J. (1998). The experience of living–dying in a nursing home: Self-reports of Black and White older adults. *Journal of the American Geriatrics Society*, *46*, 1091–1096.

Haley, W. E., LaMonde, L. A., Han, B., Burton, A. M., & Schonwetter, R. S. (2003). Predictors of depression and life satisfaction among spousal caregivers in hospice: Application of a stress process model. *Journal of Palliative Medicine*, *6*(2), 215–224.

Happ, M. B., Capezuti, E., Strumpf, N. E., Wagner, L., Cunningham, S., Evans, L. K., & Maislin, G. (2002). Advance care planning and end-of-life care for hospitalized nursing home residents. *Journal of the American Geriatrics Society*, *50*, 829–835.

Lawrence, R. H., Tennstedt, S. L., & Assmann, S. F. (1998). Quality of the caregiver–care recipient relationship: Does it offset the negative consequences of caregiving for family caregivers? *Psychology and Aging*, *13*, 150–158.

Schonwetter, R. S., Han, B., Small, B. J., Martin, B., Tope, K., & Haley, W. E. (2003). Predictors of six-month survival among patients with dementia: An evaluation of hospice Medicare guidelines. *American Journal of Hospice & Palliative Care*, *20*, 105–113.

SUPPORT Principal Investigators. (1995). A controlled trial to improve care for seriously ill hospitalized patients: The study to understand prognoses and preferences for outcomes and risks of treatments (SUPPORT). *Journal of the American Medical Association*, *274*, 1591–1598.

Engineering Psychology

Klaus-Peter Timpe, Hans-Gerhard Giesa, and Katharina Seifert

Technical University Berlin, Berlin, Germany

GLOSSARY

cognitive task analysis Any further development of traditional task analyses that yields information about the knowledge, thought processes, and goal structures that underlie observable task performance.

compatibility Consistency in the direction of the movement of relationships and operator expectations.

human–machine system A convenient abstraction of the goal-oriented cooperation between persons and technical systems to accomplish a self-determined or assigned task.

human reliability The probability of the successful performance of the human activities necessary for a human–machine system.

interface design The construction of that technical part of the machine directly interacting with the human operator on the basis of empirically grounded rules and heuristics to achieve efficient human–machine interaction as well as high system performance.

operator The human working in a human–machine system.

task analysis Any process for analyzing the behavior of a human within a human–machine system that determines the interactions between humans and equipment or machines with the underlying working conditions in mind.

usability The degree of ease with which humans can interact with a system or a piece of equipment.

The subject of engineering psychology is the optimization of the exchange of information within the human–machine system with the use of planning, analysis, design, and evaluation.

1. HISTORICAL BACKGROUND

Changes in work in all branches of industry, transportation, and services after World War II led to an increasing amount of monitoring, control, supervision, and planning activities for operators. To match the equipment to the needs and cognitive capacities of the human operators, engineers, psychologists, physiologists, and anthropologists have outlined general principles for an optimal layout of technical systems that involve people.

Since the late 1940s, the framework of specialists has focused on that topic developed as a part of human factors engineering called engineering psychology. At the beginning of the 1960s, the subject was examined in Europe, especially in England, France, and Germany. Similarly, many studies were undertaken in the Soviet Union in this new field of technical design under the

label of *ingenernaja psychologika*. In Japan, research in engineering psychology also spread quickly. Similar subjects were worked on in the United States, although socially and culturally different foci were set.

During the initial phase of engineering psychology up to around 1960, systems design research was focused on the developmental design principles of machine displays, controls, and interaction between humans and machines. However, this changed quickly due to the fast development of computing technology. By the end of the 1960s, ever more portions of the mental work of humans were already automated and new questions about the allocation of functions between humans and machines had to be answered. This led away from the "dials and knobs" optimization of the early 1950s to human–machine/computer interaction as an area of engineering psychology. Although the adjustment behavior of humans was initially modeled as an automatic controller, the field of research spread soon to investigations of the operator as a dialogue partner and interactive problem solver. System design issues requirements are currently driven, and will also be driven in the future, largely by trends in technology such as computerization, automation, knowledge, and information processing.

The new field of activity engineering psychology is closely related to other spheres of research such as human factors or human factors engineering (United States), ergonomics (United Kingdom and Germany), and cognitive ergonomics. However, differences in the core psychological matter and methodology of engineering psychology clearly separate it from these subjects in that engineering psychology focuses on information exchange in human–machine systems. The methods and tools used for analysis in the human–machine information exchanges are predominantly derived from cognitive psychology, which developed numerous paradigms to investigate fundamentally human information processes. During recent decades, an increasing trend is to incorporate and transfer methods from other scientific areas to the field of human–machine interaction studies.

2. THE HUMAN–MACHINE SYSTEM AS THE SUBJECT OF ENGINEERING PSYCHOLOGY

In general, the numerous definitions of engineering psychology can be summarized as follows: the optimization of the interchange of information within a human–machine system with the use of planning, analysis, design, and evaluation.

The concept of human–machine systems refers to a convenient abstraction of the goal-oriented cooperation between persons and technical systems to accomplish a self-determined or assigned task. This more abstract definition covers the relevant features of the diverse variations of the goal-oriented information interchange between humans and machines in different situations. The concept of a human–machine system implies that at least one person and at least one machine work together.

The general structure of a human–machine system (Fig. 1) is a control system in which an operator and/or a team make decisions to steer the technical system according to organizational basis, goal, task, and perceived feedback about the environment and process conditions. The behavior of the human–machine system is thereby determined by qualitatively different processes. On the one hand, it is shaped by human data processing based on social, psychological, and biological regularities. On the other hand, it is influenced by technical processes that obey physical laws.

The reception of information concerning process data and environmental properties takes place either directly or indirectly (e.g., a direct visual approach, an indirect instrumental approach). Information can also be entered either with appropriate devices or through language, muscle activities, and the like.

The location of these interferences and/or of the machine can be nearly arbitrary spatially. If this aspect of informational coupling is more intensely differentiated (i.e., if machines and their maintenance or control are handled by persons at different places), the term "distributed human–machine systems" may be used as well. Subtasks, responsibility, decisions, and knowledge are distributed among different human participants in different hierarchical levels so that the classical problem of task allocation arises again.
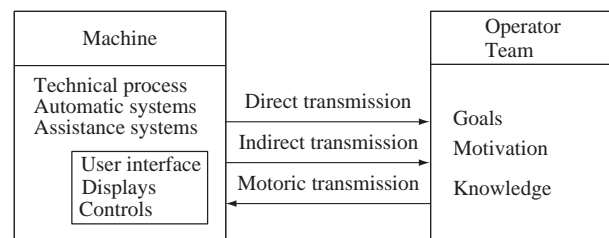

FIGURE 1 Schematic representation of information interchange in human–machine systems.

The substantial components of the technical entity "machine" in human–machine systems are indicating and control units, automated subunits, and computer-aided assistance systems (e.g., assistance or diagnostic systems) as well as the technological process controlled. Information indicators and control components are often integrated, for example, in control rooms, cockpits (for vehicles and airplanes), conning bridges, control desks, and control panels and screens.

3. HUMAN INFORMATION PROCESSING IN HUMAN–MACHINE SYSTEMS

3.1. Circulation of Information in Human–Machine Systems

Three components can be differentiated in information circulation and are represented separately for didactical reasons: the system of the initial processing of stimuli received by the receptors of an appropriate sense (perception), the memory with its substructures (information retrieval, selection, and recognition), and the system for the execution of the (motoric processing) action (Fig. 2). It is assumed that all utterances and actions by the operators are the result of information processing (e.g., comparisons, judging, decisions) in the nervous system. These processes are essentially steered by signals from the environment and by knowledge stored in memory. As shown in Fig. 2, the data are received and decoded by receptors in the eyes, in the ears, and in the haptic system (i.e., the information is perceived). The decoded information reaches the brain, where it is compared with previously stored information. The result of this comparison can be recognition or comprehension, but it can also prompt a process for behavioral decisions that result in sensomotoric activity (e.g., as steering activities). Engineering psychology considers information processing on three different levels: knowledge based, rule based, and skill based.

As a result of research in the field of general psychology, it has been shown that the evaluation of the information processed has a strong influence on
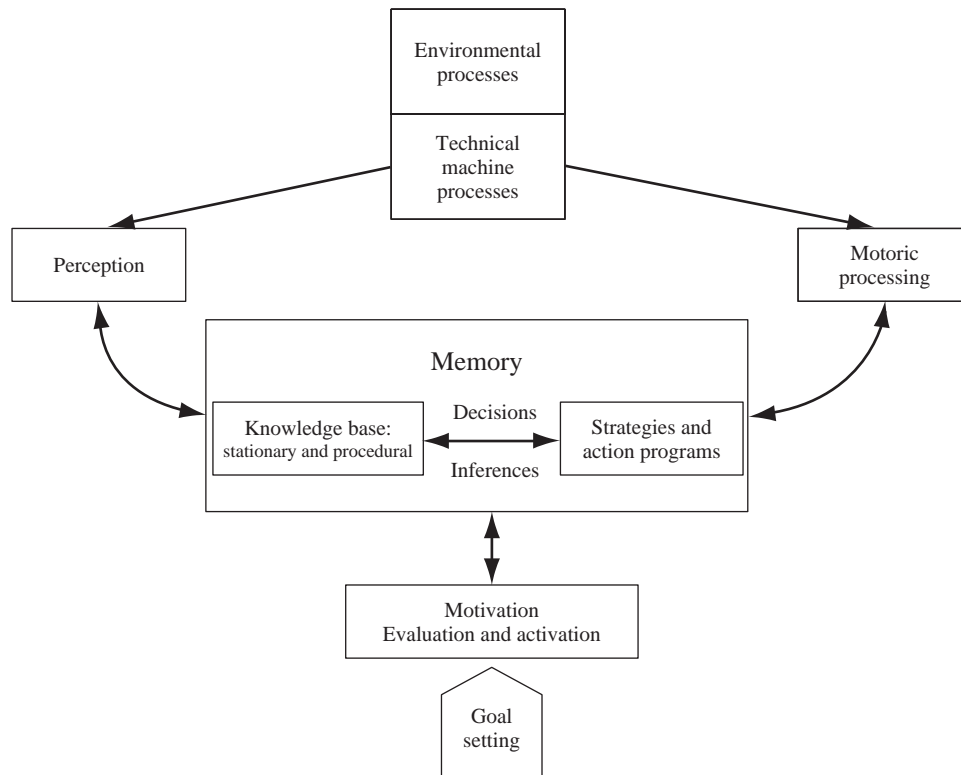

FIGURE 2 Information circulation processes in human–machine interaction (based on a structural model of human cognitive systems).

performance. This is why mental processes cannot be investigated without taking activating processes into account. This means that the interrelationship between information processing and evaluation, as well as activation processes, is an important field of research for engineering psychology (e.g., for measures of workload and situation awareness).

3.2. Goal Setting

All actions in a work context need to set a specific goal that is expected to be achieved by the actions. The goal-setting process includes an assessment of a cost-efficiency ratio as well as of a demands/effort ratio. Furthermore, the process is affected by the interpretation, motivation, and values of the operator. This can be described as an anticipation of future action and its result. The goal is to resolve problems in a definite time, to anticipate changes in the state or in the process, and/or to develop a process of execution so as to set means and define results. Goal setting can be processed at different levels of detail, depending perhaps on the amount of experience in the domain of action, on the complexity of the task, and on the degree of freedom implicated in it. Goals can be differentiated as those that are set by or taken over from another authority and those that are self-determined. However, self-determined goals are achieved with more motivation and involvement than are assigned ones. The process of goal setting is completed by means of the resolution to execute a specific action. Goal setting can be supported by the unambiguous design of displays and controls that give the operator useful hints for the current action needed.

3.3. Perception and Interface Design

Human perception is distributed over diverse channels of information intake (i.e., the senses) that are integrated into the process of perception. The senses are usually distinguished as the visual sense, auditory sense, somato-visceral sense (including the haptic sense), the olfactory and equilibrium sense, and the sense of taste. The research on perception in experimental psychology during the past 50 years or so has been predominantly directed toward the visual and auditory sensory system. In addition to these two senses, the haptic sense—especially the sense of touch and kinesthetics—has been "rediscovered" for the interface design of advanced human–machine systems (e.g., to synthetically model force feedback in distributed or virtual human–machine systems).

Psychological theories of perception differ in how they assume that information is taken in, describing the process as follows:

- A direct process that rests on the spatio-temporal invariance of the environment (ecological approach of perception)
- An indirect process involving different stages in the active construction of the percept from receptor sensations based on processing rules, the referential system, and the experience of the person involved (constructivist approach)

Current theories of perception often assume that bottom-up processes, as well as top-down processes, are involved in the various levels of perception.

Engineering psychology relies particularly on the findings of Gestalt psychology, the laws of perception, and psychophysics to offer recommendations for the design of icons for visual displays as well as of acoustic feedback. The laws of perception are related to the perceptual grouping of visual and acoustic stimuli depending on their properties. Following are a couple of examples:

- *The law of similarity:* Similar items tend to be grouped together.
- *The law of proximity:* Items with spatio-temporal closeness tend to be experienced as belonging together.

The laws of perceptual organization also explain the figure–ground disjunction and can be applied for adequate display design in human–machine systems. For example, smaller and symmetrical pieces, as well as orthogonally oriented figures, can be more easily perceived as objects.

Psychophysical findings, which are important for the design of human–machine systems (especially for information coding), are the minimal and maximal thresholds of perception for the individual senses, their discrimination thresholds, and their spatio-temporal resolution.

To support the operator in the comprehension of information, this information can be coded into features that are easy to perceive and distinguish (e.g., the color red for "hot"). Most of these kinds of codes are understandable due to knowledge about cultural conventions.

3.4. Cognition

In the future, many fields of technical design, diagnostic processes, modern management, and so forth will not remain competitive without the help of technical

aids for knowledge processing, (group) decision making, and other aspects of cognitive tasks. The ultimate aim is to support cognitive tasks as effectively and transparently as possible. Knowledge-based supporting aids and decision support systems have been developed based on findings from cognitive psychology.

3.4.1. Knowledge-Based Supporting Aids

A knowledge-based supporting aid is a technical entity for information processing that helps the operator to achieve his or her tasks in a human–machine system by taking over certain subtasks. These knowledge-based supporting aids for technical design have become an important field of engineering psychology in research and practice.

However, this knowledge must first be prepared and presented in such a way that it can be implemented in a supporting system. But this is not a matter of acquiring knowledge for knowledge's sake. Instead, it is a matter of being able to implement the relevant knowledge related to the specific task. A wide range of psychological methods (Table I) are available for the elicitation of the relevant procedural and declarative knowledge that exists, for example, as semantic networks, production rules, or schemata.

Knowledge-based diagnostic systems have been developed and applied in numerous areas, for example, for troubleshooting in medicine, mechanical engineering, or technical design. This field is witnessing a trend where engineers, engineering psychologists, and computer scientists are working together. Although the body of knowledge has been developed particularly by engineering psychologists, it has been implemented by engineers and computer scientists.

3.4.2. Decision Support Systems

Findings from the psychology of cognition should be applied to decision support or decision analysis systems. These systems contain axiomatically grounded decision procedures for managers and operators working in complex systems. Often, these working conditions are complicated by time pressure, uncertainty, information overflow, and high importance. Decision support systems in this field are frequently based theoretically on normative and descriptive perspectives on the process supported (e.g., multiattributive value theory [MAVT], multiattributive utility theory [MAUT]).

TABLE I
Methods in Engineering Psychology

Objective area	Method examples
System validation	Measurement of performance data
	Physiological measurements
	Verbal protocols
	Interviews
	Questionnaires
Cognitive task analysis	Observations
	Interviews
	Verbal reports
	Analysis of team communication
	Critical decision method
	Cognitive walkthrough
	GOMS
Knowledge elicitation	Association techniques
	Sorting techniques
	Structure formation technique
	Discussion in groups
	Grid techniques
Human reliability analysis	Analytical methods (e.g., THERP, CREAM, ATHEANA)
	Time–reliability correlations (e.g., HCR)
	Structured expert ratings (e.g., SLIM)
Usability testing	Thinking aloud method
	Constructive interaction
	Retrospective testing
	Coaching method
	Heuristic evaluation
Cognitive modeling	Computer simulations based on cognitive architectures (e.g., SOAR, ACT-R, GOMS)
Workload	Subjective measures (e.g., rating scales, psychometric techniques)
	Performance measures (e.g., primary task measures, secondary task measures)
	Physiological measures (e.g., cardiac function, muscle function, eye function, brain function)

In up-to-date systems, a team makes decisions more often. But practical design recommendations are not yet possible for the automated support of team decision making because a generally accepted concept does not yet exist. In the future, team decision support will become an important research and application domain of engineering psychology.

3.5. Motoric

Since the 1940s and 1950s, general principles for the choice and design of controls depending on working demands have already been established. Today, this information can be found in handbooks on human factors engineering or ergonomics that pertain, for example, to the optimal control/display ratio, directions of movement relationship, or coding and arrangement of controls, especially in terms of compatibility.

Compatibility is one of the most important factors to be considered in systems design. This is especially significant in stressful situations where learned habits often break down and can cause errors in action. These results from early engineering psychology can be found in all relevant compendia.

With ongoing technological development, and especially with digitalization, new concepts of human–machine interaction have been developed. Today, the research in engineering psychology is paying much attention to setting up design recommendations for speech and gesture interaction in cars and planes as well as in conventional computing. Some of these input devices are recognized as convenient and highly standardized (e.g., mice, touch screens, keyboards). They are manufactured for different practical purposes due to their widespread application in different work domains. The number of specific input devices for defined domains (e.g., endoscopic surgery) is increasing. These devices must fulfill the special demands of specific domains as well as general ergonomic criteria for evaluation. Speech input and automated image processing input (e.g., gesture input) are currently not used universally due to problems in recognition reliability, costs, and unsolved technical constraints (e.g., noise from the environment, changes in the brightness of the surroundings). The cultural aspects of the design of user interfaces, and especially of input devices, are more obvious today than they were previously due to the globalization of industrialization, and they constitute a limiting factor in standardization.

Another well-elaborated research area in engineering psychology is the modeling of the sensomotoric regulation of operators in flight and driving tasks. Formal descriptions of human sensomotoric regulation were sought at the end of the 1940s due to the requirements of military engineering. The goal of this research was to develop an analytical representation of the dynamics of a human–machine system for the choice of optimal system dimensions and to train operators on the model of the new system.

Human sensomotoric control can be described in terms of the mathematical control theory based on a structural analogy between an automatic control circuit and a human–machine system. The results of this intensive research, well known as tracking experiments, are numerous mathematically descriptive equations for the "control person's" behavior. The crossover model, the quasi-linear model, and the time-optimized model are frequently cited in this context.

Admittedly, such equations are valid only for skill-based level activities. Higher cognitive procedures, particularly the motivational and behavioral strategies of an operator, cannot be described with these equations. Today, research is concerned with developing complex operator models by means of fuzzy logic approaches, artificial neural networks, and so forth.

3.6. Attention

Attention is one of the frequently used psychological concepts relating to the main problems of automated human–machine systems design such as the following:

- Holding operators in the loop
- Maintaining the operator's positive situation awareness and availability
- Reducing the consequences of information overflow

The concept of attention refers to the human ability to focus information processing on selected events or contents over a specific amount of time. The selection of information is an important precondition for precise action. Attention lapses in interaction with a technical system may cause dangerous states or situations. But the selection of current important events is more difficult due to potentially important information and time pressure. These circumstances contribute to the mental workload of the operator. Another problem in advanced working domains is to maintain readiness for action at an adequate level of vigilance. This problem is caused by long periods of monitoring an automated system in which nothing that requires action happens or by lengthy tasks (e.g., truck driving for hours during the night). The consequences of decreasing attention or vigilance for the performance of the human–machine system are often the same: Performance falls and strain on the operator increases. Different methods have been developed to quantify actual workload levels.

4. HUMAN RELIABILITY

It is estimated that in 60 to 90% of all system failures, human actions are the main cause. In light of

these statistics, the study of human error has emerged as an important discipline of engineering psychology.

Errors in work activities are considered to be a result of inadequate training, poor working conditions, and organizational factors that are also a consequence of learning and adaptation to different situational conditions. This multifaceted interrelation means that no homogenous theory about human errors exists. Taxonomies are required for the measurement of human error. There are two main approaches for classification. The first refers to the manifestation of errors such as observably incorrect actions. These taxonomies are widely used in human reliability analysis where error frequencies must be estimated. The second approach concentrates on variations in underlying psychological mechanisms, that is, the internal causes of human error. The origin of errors resulting from incorrect action is described in various ways in the psychological literature. Both approaches are related to "active errors" that have an immediate and direct impact on the system.

To meet the requirements of the influence of organizational factors, the concept of latent errors has been established. Latent errors originate in decisions taken in the managerial and organizational spheres and may lie dormant for long periods of time.

The concept of human error is closely related to the concept of reliability. The term "reliability" has its origins in the engineering sciences and is based on the theory of probabilities. The frame of reference is the frequency of errors. On the basis of different assumptions concerning error distributions, system structures, and time intervals, the reliability of a system is calculated. The concept of reliability has been transferred from technical systems to human actions. Several attempts have been made to detect human error probabilities just like in technical systems.

Engineering psychology focuses mainly on operators and their cognitive competencies and resources. But it also must consider the organizational and situational aspects of working conditions for to evaluate the reliability of an entire human–machine system.

5. METHODS

The most important objective areas and dedicated methods used in engineering psychology are shown in Table I. They are described briefly in this section.

5.1. System Validation

System validation examines whether a machine is matched to the needs and cognitive capacities of the human operator. Validation normally takes place in an experimental setting within a controllable environment where disturbing variables can be eliminated or controlled.

During early phases of the development of a new human–machine system, single subsystems are evaluated in the laboratory. The performance of a human–machine system should be investigated under realistic conditions to achieve integrated system validation. The preparation and investigation of early prototypes have become less costly during recent years with the further dissemination and use of simulators. Technical possibilities, such as log file recordings and freezing a situation, make possible techniques of investigation that are not feasible in the real world.

A broad inventory of empirical methods from psychological research can be employed to validate new systems in laboratories or in simulators (e.g., measurement of performance data such as speed and accuracy, physiological measurements, verbal protocols, interviews, questionnaires). Workload and situation awareness are typical concepts used for system validation in engineering psychology.

5.2. Cognitive Task Analysis and Knowledge Elicitation

A new area of engineering psychology, cognitive task analysis, concentrates on cognitive aspects of human–machine interaction that are not directly observable. As a further development of traditional task analysis techniques, cognitive task analyses are performed to gain insight into the knowledge representations, thought processes, mental strategies, and goal structures that underlie task performance. The results are used for the design of human–machine interactions (e.g., task allocation, preparation of procedures); for the development of training, testing, and selection tools; for the development of expert systems; and as a basis for computer simulations.

In addition to observations, verbal statements from job experts are the main source of information for data collection. Information about cognitive processes is collected by using interview techniques and verbal reports or with the analysis of team communication. Visualization tools and sorting methods are applied for illustration. Further methods for a cognitive task analysis

are the critical decision method, cognitive walkthrough, and the GOMS architecture.

A special field of cognitive task analysis is knowledge elicitation. A wide range of psychological methods are available for elicitation of procedural and declarative knowledge necessary for the development of knowledge-based supporting aids. In addition to common methods (e.g., interview techniques, questionnaires), other methods employed include association techniques, sorting techniques, structure formation techniques, and grid techniques.

5.3. Human Reliability Analysis Techniques

Human reliability analysis is concerned with prediction of the likelihood that human performance within a human–machine system will fail. Numerous methods for human reliability analysis have been developed. Traditional HRA techniques can be divided into analytical methods, time–reliability correlations, and structured expert ratings.

Analytical methods (e.g., THERP) are based on the breakdown of a task into partial tasks or actions. Error probabilities are determined for these tasks, and actions and then integrated into an overall error probability for the analyzed task. Time–reliability correlations (e.g., THERP diagnostic model, HCR model) are used for the assessment of critical events when interpretations, diagnoses, and decisions about necessary actions must be made. They are based on a simplified correlation between available time and error probability. Structured expert ratings are used if no numerical data about the analyzed task exist. Experts can carry out an assessment with the SLIM method, which provides numerical ratings of the strength and importance of a relatively small number of factors that influence performance. Newer human reliability analysis techniques (e.g., CREAM, ATHEANA) focus on errors resulting from cognitive processes.

5.4. Usability Testing

In general, usability refers to how well individuals can interact with a product or a piece of equipment. Usability is connected with attributes such as learnability, flexibility, intuitiveness, and likability. Usability tests can be performed as a formative or summative evaluation. Formative usability tests are done to improve a system design as part of the design process.

Summative usability testing aims at assessing the overall quality of a product.

Quantitative and qualitative methods for performance measurement are used for usability testing. Primarily qualitative methods are used for formative usability testing, especially variants of the thinking aloud method (e.g., constructive interaction, retrospective testing, coaching method). Usability testing can be combined with heuristic evaluation to rely on a list of established usability principles when applied by an expert.

5.5. Computer Simulations

Models are used to analyze the attributes of a human–machine system on the basis of an abstract description of the structure and the performance of the system. In the engineering field, computer simulations have been increasingly applied in the design and evaluation of new systems. Simulations of the technical subsystem must be connected with human operator models to simulate interactive systems. In engineering psychology, models from cognitive psychology are adapted to include human behavior in the simulation of interactive systems. Cognitive architectures such as SOAR, ACT-R, and GOMS represent theories for simulating and understanding human cognition that are used for the simulation of human behavior. Detailed cognitive task analyses are needed to construct computer simulations that represent human behavior.

6. DEVELOPMENT AND PLANNING OF HUMAN–MACHINE SYSTEMS

The goal of engineering psychology is to create demands for human-oriented work design within the framework of modern information technologies. The theoretical foundations for this type of work have already been securely laid and are supported by both experimental and empirical evidence. They have been developed for successful application to systems design:

- Eliminating the negative effects of work (e.g., workload)
- High dependability
- High performance
- High usability
- General well–being and health

The main objective is to create the best possible system and to ensure that numerous components of human

information processing and task performance are included in the job requirements so that well-balanced jobs can be designed.

To achieve this goal, one must move away from the corrective systems design that dominated the initial phases of engineering psychology and move toward a preventive and flexible form of systems design. Therefore, the system's layout should be based on a dynamic task allocation rather than on a strategy of static engineering or of static human-centered automation. A system-centered automation strategy distributes functions to both the operator and the technical system so that the operator can flexibly decide whether to use automation to accomplish a task or to do it manually.

The influence of new automation strategies has moved the process of systems design from a linear–sequential design to a parallel–hierarchical and iterative procedure. Numerous process models for systems design have been developed (e.g., the V-model, the spiral model, the prototyping model). These models all incorporate empirical evaluation and phased correction into the various stages of engineering human–machine systems. Figure 3 is a rough schematic representation of this principle.

This transition from a linear systems design to a parallel one that takes the user into consideration is a major step toward reaching the goals that engineering psychology hopes to achieve by influencing the nature of modern work. The presentation of the relevant psychological theories, concepts, and empirical data in a technically usable form is crucial to the success of such a procedure. Rapid and successful developments have been made in this field, especially in aviation, automotive, and production engineering. These developments, as compared with corrective design procedures, require considerably less effort and are more economical.

7. FUTURE DEVELOPMENTS

The future development of human–machine systems is closely related to technological progress, especially to
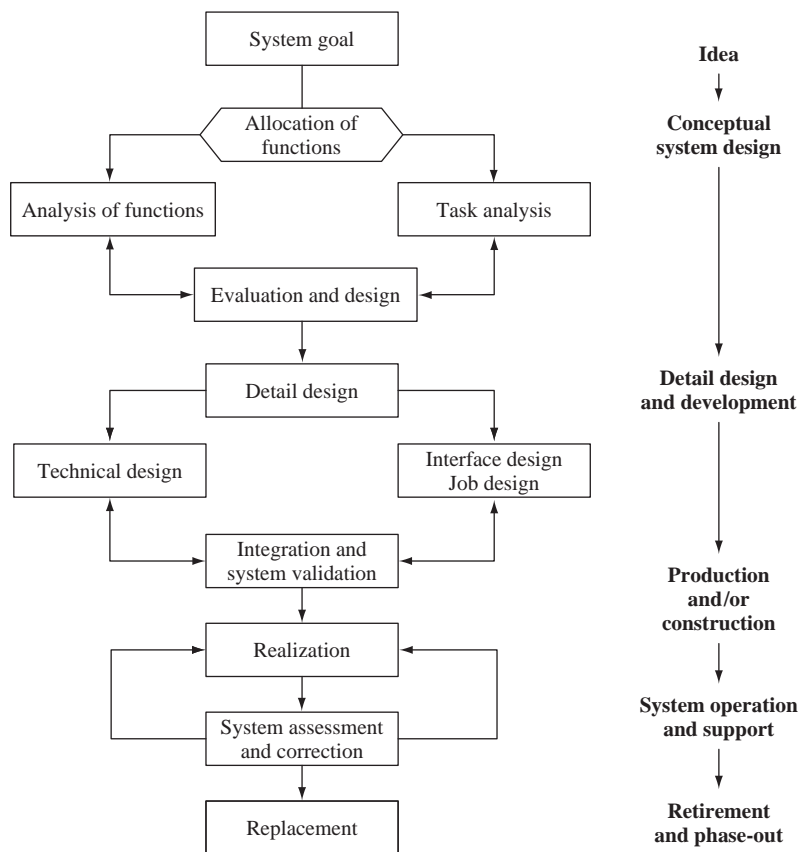

FIGURE 3 Parallel iterative systems design that includes the user.

information technology. The information exchange between human and machine will be further developed to include new interaction concepts and interaction technologies. Auto-stereoscopic perception, as well as the combined interaction of gaze, speech, and hand gestures, has been integrated into multimodal human–computer interfaces that could be applied in task environments with changing cognitive, sensomotoric, and/or attentional demands. New generations of assistance systems will observe the state of the human operator and collect safety- or task-relevant information from the environment to actively support the operator in achieving his or her aims. The development of multimodal interfaces has formulated new research questions for application and design in the field of engineering psychology, for example, about the transmission of concepts in natural communication to human–machine interaction. A special point of research in the design of interaction related to information technology is to provide access to disabled people for better living conditions.

An important concept for the supervision of complex human–machine systems is ecological interface design. This approach attempts to visualize complex processes in technical systems. The easily understandable representation of a system's states should support operators in situations where safety is critical. Other sustained work domains of engineering psychology consider the increasing amount of distributed human–machine systems throughout nearly all branches, the replacement of leftover allocation with a humanized task approach, and/or the simulation of complex human–machine systems.

See Also the Following Articles

Attention ■ Motives and Goals ■ Perception and Culture

Further Reading

Chapanis, A. (1965). *Man machine engineering*. Belmont, CA: Wadsworth.

Harris, D. (Ed.). (1997, 1999, 2001). *Engineering and cognitive ergonomics* (Vols. 1–6). Aldershot, UK: Ashgate.

Hoyos, C. G., & Zimolong, B. (Eds.). (1990). *Ingenieurpsychologie*. Göttingen, Germany: Hogrefe.

Rasmussen, J., Petersen, A. M., & Goodstein, L. P. (1994): *Cognitive systems engineering*. New York: John Wiley.

Salvendy, G. (Ed.). (1997). *Handbook of human factors and ergonomics*. New York: John Wiley.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Environmental Assessment

Rudolf H. Moos

Department of Veterans Affairs and Stanford University, Menlo Park, California, USA

Charles J. Holahan

University of Texas, Austin, Texas, USA

GLOSSARY

behavior settings Units of the environment composed of interrelated physical, organizational, and behavioral properties.

environmental assessment The conceptualization and measurement of characteristics of environments that influence human behavior.

institutional policies Rules and regulations that organize everyday social interaction and the overall functioning of a setting.

institutional structure Objective characteristics of organizations such as their size, ownership, and centralization of authority.

physical and architectural features Objective characteristics of the natural and built environment, including the landscape, cities, and buildings.

social climate The social characteristics and expectations of a setting such as its level of support, task orientation, and structure.

suprapersonal factors The aggregate characteristics of the individuals in a setting, including their preferences, abilities, and behavior patterns.

An integrated perspective on human environments is essential for one of the central tasks of applied psychology, that is, to help create optimal and sustainable conditions that enhance individuals' quality of life. At its most basic level, environmental assessment involves the conceptualization and measurement of characteristics of environments that influence human behavior. More broadly, environmental assessment encompasses the linkages among aspects of the environment as well as the processes by which environmental factors influence human behavior.

1. MAJOR DOMAINS OF ENVIRONMENTAL FEATURES

The main characteristics of environments can be conceptualized in four interrelated domains: physical and architectural features, institutional structure and policies, suprapersonal factors, and social climate.

1.1. Physical and Architectural Features

The assessment of physical and architectural features typically involves direct observation. Four main sets of dimensions have been identified: (a) physical amenities that improve convenience and comfort, (b) physical features that provide support for individuals, (c) size

allowances for personal activities, and (d) physical integration with the surrounding community as assessed, for example, by the extent to which the community and its resources are convenient and accessible.

A related area of research focuses on landscapes and natural environments. The assessment of natural settings involves ratings of perceived environmental quality as well as cognitive mapping techniques that tap people's environmental comprehension. Natural settings have been described in terms of physical attributes (e.g., the presence and extent of hills, water, trees, and grass) and aesthetic qualities (e.g., beauty, harmony, complexity). Cognitive mapping, or the study of people's internal representation of the external world, is especially helpful for understanding the value of environmental cues for orientation and "way finding."

1.2. Institutional Structure and Policies

There are four sets of characteristics in this domain. The first set encompasses structural factors such as ownership, size, and centralization. The second set of dimensions focuses on the balance between individual freedom and institutional order as measured by how much participants individualize their activities and routines and how much they are involved in managing the environment and influencing its policies. The third set of dimensions involves behavioral requirements such as expectations for individuals' functioning and acceptance of problem behavior. The fourth set of dimensions involves the provision of services and activities such as health services, daily living assistance, and social activities. These characteristics can be measured by direct observation or by ratings of participants or outside experts.

1.3. Suprapersonal Factors

When individuals come together in a social group, they bring with them values, norms, and abilities. Because of selective mechanisms, groups draw their members from the general population in a nonrandom manner and produce distinctive blends of these individual characteristics. The aggregate of the individuals in a setting, or the suprapersonal environment, in part defines the subculture that develops in a group.

In the vocational or work domain, Holland showed that occupations can be categorized into six groups (realistic, investigative, artistic, social, enterprising, and conventional), each of which represents a different personality type. Because people in a vocational group tend to have similar personalities, they are likely to create characteristic interpersonal environments. Thus, there are six types of environments that correspond to the six personality types. In this way, an average background characteristic of a group (i.e., its dominant vocational preference) creates a characteristic environment with unique demands, rewards, and opportunities.

1.4. Social Climate

The social climate is the "personality" of a setting or environment such as a workplace, a classroom or school, a social group, or a neighborhood. Three sets of dimensions characterize the social climate of diverse settings. First, relationship dimensions measure how involved, supportive, and expressive people are in a setting. Second, personal growth dimensions tap ways in which an environment encourages personal growth such as by emphasizing autonomy, achievement or task orientation, competition, intellectual orientation, and moral–religious orientation. Third, system maintenance and change dimensions measure the settings' organization, clarity, control, and responsiveness to change. In general, social climate has been measured by the shared perceptions of the individuals in an environment and by outside observers' ratings.

2. LINKAGES AMONG ENVIRONMENTAL FEATURES

The four domains of environmental features are interconnected. Physical and architectural features, institutional structure and policies, and suprapersonal factors shape each other and combine to create behavior settings with particular social climates. In turn, behavior settings and the social climate influence individuals' health and well-being both directly and indirectly through the other environmental domains.

2.1. Behavior Settings

Behavior settings, which are composed of groups of individuals behaving together (e.g., working on a shared task, participating in a social group, learning in a classroom), have four essential attributes. First, behavior settings have one or more standing patterns of behavior such as sitting, listening, talking, and note taking during a union meeting. Second, behavior settings involve a

physical milieu that may include both natural and constructed features of the environment. Third, these settings have shared policies and expectations about acceptable behavior. Fourth, the physical milieu and standing patterns of behavior are interdependent so that, for example, chairs in a classroom typically face the teacher for listening and observing. The physical milieu and policies help to shape the standing patterns of behavior in a setting. These three characteristics, in turn, tend to create a particular social climate.

2.2. Physical Features and Social Climate

Physical features that improve convenience and comfort, such as physical amenities and social–recreational aids, are associated with a more cohesive and well-organized social climate. Residential settings with more supportive physical features, such as prosthetic aids and safety features, tend to develop a social climate oriented toward independence. In contrast, crowding and lack of adequate space are associated with a lack of cohesion and goal direction as well as with ambiguity and disorganization.

2.3. Institutional Structure and Policies and Social Climate

Ownership and size are related to social climate. Policies in nonprofit facilities are more accepting of problem behavior and more likely to promote residents' personal control. Accordingly, residents in nonprofit facilities see them as more cohesive, self-directed, and better organized than do residents in proprietary facilities. With some exceptions, large size is related to a less cohesive and less well-organized social climate with somewhat more conflict. Policies that enable individuals to determine their daily routine and to influence the operation of the setting, and those that have more provisions for privacy, tend to promote a cohesive social climate oriented toward autonomy and independence. When the rules and regulations in a setting are clear, the social climate is likely to be more cohesive and self-directed.

2.4. Suprapersonal Factors and Social Climate

The modal personal characteristics of individuals in various occupations are associated with the work climates they create. For example, individuals in realistic occupations, such as engineers and air traffic controllers, tend to see their work groups as high in control and lacking in cohesion, autonomy, and innovation. In contrast, employees in social-type work settings, such as teachers and social workers, tend to report more emphasis on supportive relationships, autonomy, clarity, and innovation.

Individuals' abilities, personal resources, and typical patterns of behavior help to determine the social climates in residential facilities. Facilities with more women and more socially privileged individuals tend to develop more cohesive, self-directed, and well-organized social climates. Facilities with more older and mentally impaired individuals are likely to be less supportive and goal-directed as well as higher on control.

3. INFLUENCE OF ENVIRONMENTAL FACTORS

Each of the four sets of environmental factors has an impact on individuals' health and well-being. In addition, part of the influence of environmental factors depends on the personal competence, orientation, and preferences of the individuals in the environment. In turn, these individuals select and shape the environment they inhabit.

3.1. Environmental Impacts

Physical and architectural features, institutional structure and policies, and suprapersonal factors tend to exert their effects through specific behavior settings and the social climate they create. Most basically, social systems tend to maintain and accentuate individual qualities that are congruent with their dominant aspects. Accordingly, youngsters in families that value independence, achievement, and intellectual and recreational pursuits are likely to develop more self-confidence and social competence. A workplace that values independent decision making, complex and challenging tasks, and high-performance expectations promotes intellectual flexibility and self-efficacy. More cohesive and socially integrated neighborhoods characterized by adult guardianship and monitoring tend to promote responsible parenting and youth behavior.

Some emphasis on each of the three domains enhances positive emotional and behavioral development, but too much focus on any one domain can lead

to distress and dysfunction. Highly achievement-oriented and structured families can create anxiety and erode youngsters' self-confidence. Similarly, high work demands can lessen employees' job morale and elicit depression and physical symptoms. Difficult and competitive learning environments encourage cognitive growth, but they also can erode students' self-confidence and produce high absenteeism and dropout rates.

3.2. Person–Environment Matching Models

Part of the influence of contextual factors depends on the personal competence, orientation, and preferences of the individuals who experience them. The competence–demand model of person–environment congruence states that high environmental demands should have more positive consequences for individuals who are functioning well than for those who are functioning poorly, whereas environmental resources should have more benefit for less competent individuals. Specifically, a self-directed setting that has high performance expectations and relatively little structure should have the most benefit for functionally able individuals. In contrast, less competent individuals need more support and structure and often find high performance expectations and self-direction to be disruptive. As individuals' cognitive and psychosocial skills improve, they should be able to cope with more demanding and less structured environments.

The conceptual-level model describes people in terms of their cognitive complexity and describes environments in terms of their structure. It posits that individuals at a high conceptual level are able to organize their own environment, whereas individuals at a lower conceptual level need the stabilizing influence of a well-structured setting. The model also proposes that individuals with less cognitive complexity are more affected by changes in structure than are those at a more mature level.

3.3. Transactional Models

Most of the research in this area reflects a social causation perspective in which the environment is seen as a causal factor that shapes individuals' development and maturation. However, there also is a social selection process by which individuals select and shape existing environments and create new environments. According to behavior setting theory, four transactional processes explain how ongoing interactions among participants and other aspects of behavior settings produce a stable patterned set of behaviors and environment–behavior congruence. These processes involve goal circuits (i.e., how participants set goals and achieve them), program circuits (i.e., how participants control and organize the program or agenda of a setting), deviation-countering circuits (i.e., the ways in which unwanted modifications or problems in the setting are altered or corrected), and vetoing circuits (i.e., the ways in which deviant components of a setting are eliminated).

More generally, there are four main processes by which individuals and environments influence each other. In a passive process, individuals find themselves or are placed in an environment that shapes their behavior, for example, when a child is placed in a group care home. In an evocative process, individuals select specific responses that become part of their environment, for example, when a youngster with a sunny disposition elicits a supportive context. In an active process, individuals select social contexts that maintain and accentuate their dispositions or skills, for example, when a youngster joins a soccer team or a debating club. Finally, in a transcending process, individuals build entirely new physical and social contexts, for example, when individuals bereaved and maimed in a cataclysmic disaster rebuild their community.

4. CONCLUSION

Environmental assessment provides conceptual and empirical tools that make environmental determinants of human behavior accessible to psychological inquiry. Researchers have identified four main domains of environmental features, assessed specific dimensions in these domains, and examined their interconnections. Conceptual models of how contexts influence individuals, of person–environment matching, and of how individuals select and create contexts have led to a better understanding of how the interplay of environmental and person factors shapes individuals' health and well-being. Environmental assessment complements the traditional person-centered focus in psychology by providing a more holistic and integrated understanding of human behavior as it occurs in real-world settings.

See Also the Following Articles

Environmental Design and Planning, Public Participation in ■ Person–Environment Fit

Further Reading

Barker, R. G., & Schoggen, P. (1973). *Qualities of community life*. San Francisco: Jossey–Bass.

Bechtel, R. B., & Churchman, A. (Eds.) (2002). *Handbook of environmental psychology*. New York: John Wiley.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Friedman, S. L., & Wachs, T. D. (Eds.) (1999). *Measuring environment across the life span: Emerging methods and concepts*. Washington, DC: American Psychological Association.

Golledge, R. G. (Ed.) (1999). *Wayfinding behavior: Cognitive mapping and other spatial processes*. Baltimore, MD: Johns Hopkins University Press.

Holland, J. (1992). *Making vocational choices: A theory of vocational personalities and work environments* (2nd ed.). Odessa, FL: Psychological Assessment Resources.

Moos, R., & Lemke, S. (1994). *Group residences for older adults: Physical features, policies, and social climate*. New York: Oxford University Press.

Walsh, W. B., & Craik, K. H. & Price, R. H. (Eds.) (2000). *Person–environment psychology: New directions and perspectives* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Environmental Design and Planning, Public Participation in

Arza Churchman

Technion–Israel Institute of Technology, Haifa, Israel

Elisheva Sadan

Hebrew University of Jerusalem, Jerusalem, Israel

GLOSSARY

civil society The part of social life that lies beyond the immediate reach of the state and that must exist for a democratic state to flower; it is the society of households, family networks, civic and religious organizations, and communities that are bound to each other primarily by shared histories, collective memories, and cultural norms of reciprocity.

community A group or network of people who share something in common or interact with each other; it is a place where solidarity, participation, and coherence are found.

distributive justice The equitable distribution of resources and nuisances among groups within the society differentiated by factors such as power, social class, ethnicity, and so forth.

empowerment The process by which people achieve more control (perceived and actual) over their lives, destiny, and environment; it is the process of overcoming powerlessness and its outcomes (e.g., despair, marginality, estrangement).

environment An all-encompassing term that includes all aspects of the world—the physical, ecological, social, economic, cultural, political, institutional, technological, and individual.

participation/involvement Situations in which individuals (or groups), who are not elected or appointed governmental officials and who are not professionals working for such officials, voluntarily take part in decision-making processes to influence these processes, change them, or improve them; participation is the activity of such individuals, whereas involvement is the activity of the officials and professionals who initiate such a process.

participatory democracy A democratic system in which citizens have opportunities for taking part in governmental decision-making processes relevant to their lives on a regular basis and not just in elections once every 4 years.

participatory governance The transfer of authority and responsibility from those who hold power by virtue of law, contract, or organizational role to those not so empowered; this is a step above participatory democracy because it makes explicit the authority and responsibility given to the participants within the governmental procedure.

procedural justice A system in which the decisions as to the distribution of resources and nuisances are made in an equitable manner.

social capital The bonds of the community that enrich people's lives in a variety of ways; it includes resources that inhere in social relationships such as mutual trust, a sense of reciprocal obligations, and civic participation aimed at benefiting the group or community as a whole.

sustainability The goal for development that recognizes the interwoven nature of economic, social, and ecological–environmental factors and that strives for intergroup and intergenerational equity.

sustainable development The process of balancing three development processes: economic, community, and ecological; it is development that meets the needs of the current generation without compromising the ability of future generations to meet their own needs.

Participation relates to people taking part in public decision-making processes or private initiatives with public implications. The recent movement in planning and design toward more participatory and empowering decision-making processes is a significant step in the direction of accommodating the needs and preferences of different groups of people. Introducing individuals into the decision-making process forces politicians, planners, and architects to consider the micro level and to meet the people, see their faces, and hear their voices. The interest in such processes within environmental psychology forces one to consider the macro level and to understand the decision-making system and the constraints and pressures within which planners operate. This article addresses various issues that reflect the complexity of the participatory process, the discussion of which also illustrates the interdependence of these issues. Finally, the article looks at the role of environmental psychology in the public participation arena.

1. INTRODUCTION

- "I don't need to involve the public. I know what's good for them."
- "They didn't like the Eiffel Tower at first either."
- "Of course I'm in favor of public participation ... *but* ... the public isn't ready, or they don't understand, or it takes too long, or it costs too much."
- "Who are you to claim that you represent the public?"
- "Why aren't there more people here?"
- "We've been working on this plan for you for 3 years, and now when it's ready you're against it!"
- "You're only interested in your own backyard and your own selfish interests, while I'm concerned with the larger picture and the common good."

These are quotes of statements made by elected officials and professionals with regard to the principle of public participation in decision making. They illustrate some of the problems encountered by those who are interested in facilitating or promoting the principles of public participation, whether they are professionals, elected politicians, or members of the public themselves.

Participation, in the sense that the term is used in this article, relates to people taking part in public decision-making processes or private initiatives with public implications, not those occurring within an organization or a group of people. Arguments over what should and should not be considered participation illustrate both the ambiguity and the political nature of the concept. For example, there are those who talk about participation in terms of voting in elections, and others who see it in terms of taking part in groups of various kinds. It is clearly a value-laden concept, which for some represents the very essence of democracy, a basic human right, and an ethical good in and of itself. Others value participation as a pragmatic means to various ends, some instrumental and some strategic. Participation is seen as having the potential for major benefits for the individual, for the community, and for society as a whole; as enabling decisions that are better able to fit the needs of the people; and as making support for policies and programs more probable.

Because participation has implications for the nature of the distribution of power among individuals and groups, it arouses tensions and emotions, and this makes it difficult to relate to participation dispassionately. On the whole, politicians, officials, and professionals see it as a threat to their power and as a bother. They argue that ordinary people are not capable of being part of the decision-making process due to the complexity of the issues at hand. They also argue that they see the larger picture, and make their decisions "objectively," for the greater good. Residents, on the other hand, are "subjective" and concerned only with their own private good.

The concept of public participation appears in the writings of philosophers, theoreticians, and practitioners in the fields of political science, sociology, planning, architecture, landscape architecture, community psychology, community social work, and environmental psychology. Because one of the basic tenets of environmental psychology is the importance of identifying the context of a phenomenon for understanding it, the next few paragraphs are devoted to identifying the historical, political, intellectual/theoretical, practical/professional, cultural, and psychological contexts of public participation.

Participation is a multifaceted and multidisciplinary project, with theoreticians mainly to be found in political science and planning. The current discourse surrounding participation is rooted in the theories of Habermas and Giddens that have been further developed in the planning context by Healey and Friedmann. Freire has had an important influence on the thinking in this area as well. Recently, Putnam's theory of social capital has also found its way into the discussion of the role of the civil society and of participation in current democracies. There is also a body of knowledge about participation within community psychology, both theoretical and empirical. On the other hand, one can find professionals actively involving others within the fields of community organization, architecture, planning, and environment–behavior studies, including environmental psychologists. There are private design offices that specialize in participatory work (particularly in the United States), and there is growing interest in participatory research and participatory evaluation.

Attempts at involvement and at participation appear within rather varied contexts with different professions involved, depending on the context and often with different motivations and methods, with regard to environmental quality issues, health, housing, neighborhoods, schools, public spaces, and community development in general. United Nations (UN) agencies and documents, such as the Convention on the Rights of the Child, Agenda 21, Habitat 2, and the Healthy Cities Program, strongly recommend (and sometimes require) involvement of the public. In the European Union, the Aarhus Convention emphasizes public participation in decision making as a necessary component of governmental activity in the area of environmental policy. An interesting phenomenon during recent years is the promotion, or even requirement, of public participation by international agencies, such as the World Bank, operating in developing countries. Public participation appears to be the recent offspring of community development, which was a UN strategy for "developing" those that were then called Third World countries. That approach to development occasioned much intellectual and academic criticism due to its poor outcomes in those countries, and it would seem that a more empowering involvement approach is the new consensus. Because a democratic system is the basic necessity for the right to participation in that it requires inclusiveness, transparency, and accountability on the part of those in positions of power, attempts to apply its principles in nondemocratic systems are limited to topics that are perceived as not threatening the system as a whole.

The development over time in the approach to participation can be seen in the changes that have occurred in the language used to discuss the ideological bases for it—from democracy, participatory democracy, and power to justice (both distributional and procedural), sustainability, empowerment, social capital, and participatory governance. These changes in language reflect changes both in the theoretical underpinnings of the concept and in the degree to which it is embedded within public decision-making processes. These reflect the change from the 1960s "modern" and positivistic view of the professional and the needs-based participation process to the "postmodern" times of reflective practitioners and the Foucauldian thinking that knowledge is not where (and what) it was traditionally believed to be and that power must be shared or at least viewed with suspicion.

In many countries and cities, resident involvement now appears in the language of master plans as a basic requirement. For example, this is the case in Australia, virtually all of Europe (but particularly Great Britain and The Netherlands), Canada, the United States, and Israel. In these and other cities and countries (e.g., Brazil, Mexico, Venezuela, India, South Africa), there are examples of participatory processes of greater or lesser scope. Participatory processes have been conducted with adults and children, with men and women, with poor people and middle-class people, and with people with and without formal education.

Furthermore, "in the beginning" there was involvement, and even though the term used might have been participation, most of the discourse in the planning and architecture fields was about the activities of the professionals and how they should or could involve others. Whereas once the discourse of public participation was dominated by professionals who were "doing it" and the voice of the participants was never heard, there is much more recognition now that professionals are not responsible for the participation but rather are responsible for effectively involving people in processes that, until recently, professionals have controlled on their own. Thus, there is growing awareness today that the participants, their motivation, and their concerns should be more in focus.

To use an image suggested by Kidder and Fine in 1987, the lens used in planning is to a large extent a zooming-out one, whereas the lens used in environmental psychology is a zooming-in one. Public participation was usually discussed through the zooming-out lens, and this article zooms in on it. The movement in planning and design toward more participatory and

empowering decision-making processes is a significant step in the direction of accommodating the needs and preferences of various groups of people. Introducing individuals into the decision-making process forces politicians, planners, and architects to consider the micro level and to meet the people, see their faces, and hear their voices. The interest within environmental psychology in such processes forces one to consider the macro level and to understand the decision-making system and the constraints and pressures within which planners operate.

On a very basic level, many of the assumptions of environmental psychology should lead to the conclusion that participation is the only way, or the best way, in which to achieve the desired "fit" between the environment and its users. These assumptions include (a) that the physical environment has implications for people's lives; (b) that people are different and have different needs and, therefore, that one cannot specify, identify, or posit one universal model of the person–environment relationship; (c) that people are active and not passive in the sense that they interpret, evaluate, and use their environment in ways that they desire or are able; (d) that freedom of choice is an important element in people's behavior; (e) that aggregates (or groups) that have some needs and characteristics in common can be identified, for example, by age, gender, health status, socioeconomic status, and cultural background; and (f) that the purpose of the environment is to afford opportunities for achieving each individual's own subjective definition of quality of life. The fact that participation is not yet seen as a sine qua non of environmental psychology is an unfortunate circumstance.

2. ISSUES BEDEVILING PARTICIPATION

This section addresses various issues that reflect the complexity of the participatory process, the discussion of which also illustrates the interdependence of these issues.

2.1. Why Involve Others and Why Participate Oneself?

If participation is seen as an end in itself, perhaps the question of what it is good for is unnecessary. However, if participation is seen only (or also) as a means to other ends, major questions arise as to whether it indeed can

achieve (or does achieve) these ends. Related to this question is the distinction between the process of involvement and/or participation and the products and/or results of this process. For example, one could look at how many people participated or who they were. One could look at whether participants improved their skills and abilities, whether there was individual or community empowerment, or whether there was a change in the level of trust between the involvers and the participants. Then, one could look at the product of the decision-making process. Is it a "better" one (and by whose criteria)? Did it lead to an improvement in the quality of the environment? There are a number of issues pertinent to this distinction. Which is more important, process or product? For whom is it more important, involvers or participants? Under what circumstances? How does one measure and evaluate the elements of the process or the results (e.g., quantitatively or qualitatively, by effectiveness or by efficiency, subjectively or objectively)? How does one determine that particular results are directly related to the process?

Another question that arises is whether the adherents of participation are too ambitious or unrealistic in their expectations and goals. Are their assumptions, theories, or hypotheses ones that can be achieved? Just as it has been argued that architects and planners should not make deterministic statements that promise specific results from their plans, it might not be appropriate to promise too much from a participatory process and raise expectations that might not be possible to achieve. When this happens, it may doom the project of participation to disappointment and rejection, both by the participants and by the involvers. It would seem that what is needed is (a) a careful analysis of what is possible to achieve under what circumstances, (b) a serious effort at implementing the process in a manner that has the potential for achieving those results, and (c) research and evaluation that will enable one to learn what does and does not work. On the other hand, the participants must understand and accept the fact that they might not be able to achieve all that they wish to achieve, that compromise might be necessary, and that consideration for the views and needs of others is an essential part of the process.

If one focuses the lens on the participants themselves, one can ask whether it is enough if the results are mainly on the psychological level, which is also the hardest level to demonstrate. If results happen on the behavioral level, such as attendance at meetings, how does one know that that there are also results, such as personal empowerment, on the psychological level? To

be sure that the claims for the results of participation are valid, one must develop tools for evaluation that include indicators and "signs" of the psychological, educational, and environmental impacts or outputs of participation.

To paraphrase the words of Schorr, to be more assertive about what one knows, one must first be more systematic in assembling and trying to understand what one knows. One will not disappoint if one bases his or her recommendations on "what works" in its specific context. If one is systematic about documenting participatory processes—and the works of Fung, Chawla, Couto, Horelli, Sanoff, Schneekloth, Schorr, and others show that this is the trend now—one will be able to say more confidently, with accumulating evidence, what results are caused by the process. The accumulation of research results pointing in the same direction helps to build confidence in the hypotheses.

2.2. Why Is Involving Others So Difficult?

Many of those who discuss participation or practice involvement are unaware of the complexity of the concept and of the variety of ways in which it can be implemented. A one-dimensional approach to involvement, such as public hearings as the sole medium, ignores the various possible goals for involvement and participation and the differences among the participants themselves. This article has already noted that some see participation as an end in itself, whereas others see it only (or also) as a means to other ends. Some of these other goals focus on the participants and on the manner in which they can achieve personal and group change through participation. Other goals focus on the decision-making process and the manner in which participation can change this process and its results. Different professions focus on different goals, although it is important that all professions are cognizant of all such goals. For example, environmental psychologists and architects and planners may focus mainly on the goal of achieving environments that better fit the needs and preferences of the various users, whereas community psychologists may focus more on the manner in which the participatory process leads to more community cohesiveness or personal skills. However, if the former ignore the fact that participants have personal and group motivations that may drive their willingness to participate, and if the latter ignore the fact that participants are also concerned with other kinds of tangible results, both

groups may fail to achieve their goals. Until now, those working in this area have not really examined systematically whether individual goals are realized. One reason is that they are too busy doing participation to do this. Another reason is that one faces the dilemma of how to "prove" that there is a cause-and-effect relationship. To gain the trust of the people, one must be honest about the constraints within which the process will be operating and be careful not to raise expectations beyond what is likely to be possible.

Part of the problem stems from a lack of understanding of the complexity of participatory processes, that is, that there is not one way that is appropriate for every context or situation. Furthermore, many lack knowledge as to how to work with people; working with people is perceived as something that anyone can do easily. If one really wishes to involve people, one must make sure that he or she has the necessary skills of listening, knowing how to talk to lay people, honesty, perseverance, and understanding of people's needs and interests. It must be a process in which people will be able to participate and will want to participate and that will allow people to make decisions on questions that are meaningful to them. One must demystify the professional process, avoid jargon, and exemplify the essence of better communication, namely simplicity, clarity, and adaptability.

2.3. Who Is the Public?

The issue of how to define the public concerns the involvers, who wish to have some assurance that those participating "represent" the public at large. Considering the fact that every "public" contains many different groups with diverse and sometimes conflicting interests, this is not a trivial demand. It raises the issue of the possible conflict between what is called the common good and the self-interest of individuals or groups. There are a number of answers to these questions. One is that there is no objectively defined common good; it is always a value judgment as to which interests to prefer over others. Another answer is that self-interest is not necessarily "selfish-interest." Recognizing self-interest can be seen as helping to advance distributive justice. The more groups and interests participate in the deliberations, the more differences can be clarified and conflicts can be addressed. Still another answer is that those who participate can be viewed as informal representatives of those who do not attend, particularly if serious efforts

have been made to inform all of those relevant to the topic of the possibilities for participation.

Although the participants themselves might not be particularly concerned about this issue of representativeness, it does affect the degree of legitimacy of their claim to attention and their ability to influence the process. Some address the issue in quantitative terms (e.g., how many attend a particular event or process), whereas others do so in more qualitative terms (e.g., how varied or how serious those who participate are).

In a situation where there are professionals or decision makers who initiate the participatory process, it is incumbent on them to make sure that there are no initial stumbling blocks to the participation of as many people as possible (e.g., meetings at inconvenient times or in inconvenient places, notices that people cannot read or understand). In some situations, this requires preparatory work within the neighborhood or surroundings so as to raise people's awareness of and interest in the problems being addressed. The process itself must be a welcoming one that speaks to people in ways that they understand, that does not condescend or mislead, and that enables the empowerment of the participants. People must be provided with the information relevant to the questions being addressed, both those questions considered to be relevant by the involvers and those considered to be relevant by the participants.

On the other hand, those who are to be affected by the decisions taken must be willing to be part of the process, to understand that it is not only their right but also their responsibility to take on the role of participants. It is to their advantage to do so in terms of the personal and social capital that they can accrue and in terms of the manner in which they can work toward changing both the system and its decisions to better reflect their needs. However, such participation requires skills that not everyone has, and it may be necessary to help participants to acquire those skills and, in so doing, enhance their self-confidence and self-efficacy. Furthermore, it must be accepted that there are members of the public who are not interested in the topics or in the process and that this also is their right. Some people have such serious problems of various types that they have no time or energy to devote to other interests.

In other circumstances, there are cases where it is the public (or parts of it) whose members demand to be part of the decision-making process. They too have a duty to create a democratic process within their group and to work toward the inclusion of people with different characteristics and different ideas. In this way, they

will have a stronger case when they are confronted with the inevitable questions of "Who are you?," "Why should we listen to you?," and "Who chose you?"

2.4. Whose Values Are Primary?

Very often in participatory processes, there is conflict between the values of the professionals and decision makers and those of the public. The former, particularly in the environmental quality area where some of the issues seem to be more clear-cut, believe that they cannot leave the decisions up to the "ignorant" public. With these environmental issues, there is an agenda that they want to push where there are ostensibly more objective kinds of issues with which to deal, and the professionals are convinced that they have the answers. The fact that people are likely to mistrust the statements and pronouncements of politicians and professionals makes their task much more difficult. One way in which to gain (or regain) trust is to conduct a sincere participatory process.

However, there is a general value that should be recognized—one that agrees to disagree and that understands that the dialogue itself is important. The involvement/participation process should acknowledge the conflicts among various groups and, instead of denying them and deciding that the politicians and professionals will have their way each time, should encourage the democratic participation of others involved in the decision-making process and enable the expression of the differences of opinion that exist in the social arena. This is the whole purpose of what Couto called "making democracy work better." Rather than a question of whose values are primary, it is a question of representing the diversity of values involved in the issue. Sandercock called this working for a society in which differences can flourish based on principles of social justice, of multiple citizenships, of heterogeneous publics, and of coalitions building bridges across differences.

2.5. How Does One Involve Others?

Currently, there are a number of excellent sources that present many different techniques for involving the public, some more detailed than others but all extremely valuable for illustrating the variety of techniques available and thereby expanding the ways in which participation is viewed. These include, but are not limited to, the works of Chambers, Driskell, Horelli, Sanoff, Sarkissian and colleagues, and Wates. Many of

these publications are useful not only in presenting the techniques but also in spelling out the pitfalls to avoid in involvement processes. In addition, there is growing use of geographic information system (GIS) and Internet techniques, although limitations on their use with groups whose members might not have access to computers or the Internet must be taken into account, just as the use of plans or models raises questions as to whether or not nonprofessionals understand them.

3. THE ROLE OF ENVIRONMENTAL PSYCHOLOGY IN THE PUBLIC PARTICIPATION ARENA

The contribution of environmental psychology to the discourse and practice of public participation in decision making lies in a number of directions and stems from theoretical developments and practical experience. On the theoretical level, environmental psychology contributes a number of basic tenets, including (a) the definition of environment as an all-encompassing term, (b) the emphasis on context and on the relationship between sociophysical aspects of the environment and human behavior and attitudes, (c) the knowledge that both physical and social conditions can be acted on rather than accepted fatalistically as fixed and unchangeable, and (d) the recognition of the differences in behavior and attitudes among different groups of people and the importance of this for the basic philosophy of participation. The research is based on notions that are also basic to participation—that there is no one best solution to a design problem, that expert decisions are not necessarily better than lay decisions, and that a design or planning task can be made transparent. Environmental psychologists believe in the principle that the environment works better if it is designed and planned in such a way that it fits the needs of the people affected by its changes. Environmental psychologists involved in participation add to this the condition that these people are actively involved in its creation and management. Their role is to focus the spotlight on the people themselves—on participation rather than involvement and on the people and their lives rather than on abstract concepts.

In more specific terms, the research contributes an understanding of, for example, how people perceive environmental problems, what the determinants and predictors of environmental concern are, attitudes toward risk and perceived lack of control, and the relative importance of different environments for different groups. The latter point is particularly relevant given the finding by many that people are more concerned with immediate and local issues than with wider and more abstract ones. Environmental psychologists have experience in research that focuses on real-world complex issues, the kind of research that is essential when it comes to researching participatory processes. This focus on research, if applied to the phenomena of participation, will enable the examination of what does and does not work and thereby further the paradigmatic shift that participatory processes entail. However, this will require that more environmental psychologists make that paradigmatic shift themselves, both in their research and in their work within design, planning, and policy decision-making processes.

See Also the Following Articles

Decision Making ■ Environmental Assessment ■ Residential Preferences and Attachment across the Lifespan ■ Restorative Environments

Further Reading

Braydon-Miller, M., & Tolman, D. (1997). Transforming psychology: Interpretive and participatory research methods [special issue]. *Journal of Social Issues*, *52*(4).

Chambers, R. (2002). *Participatory workshops*. London: Earthscan.

Chawla, L. (Ed.). (2002). *Growing up in an urbanising world*. London: Earthscan.

Churchman, A. (2002). Environmental psychology and urban planning: Where can the twain meet? In R. Bechtel, & A. Churchman, (Eds.), *Handbook of environmental psychology* (pp. 191–200). New York: John Wiley.

Couto, R. A. (1999). *Making democracy work better: Mediating structures, social capital, and the democratic prospect*. Chapel Hill: University of North Carolina Press.

Driskell, D. (2002). *Creating better cities with children and youth: A manual for participation*. London: Earthscan and UNESCO.

Freire, P. (1985). *The politics of education*. South Hadley, MA: Bergin & Garvey.

Friedmann, J. (1992). *Empowerment: The politics of alternative development*. Cambridge, UK: Blackwell.

Fung, A., & Wright, E. (2003). *Deepening democracy. Institutional innovations in empowered participatory governance*. London: Verso.

Healey, P. (1997). *Collaborative planning*. London: Macmillan.

Horelli, L. (2002). A methodology of participatory planning. In R. Bechtel, & A. Churchman (Eds.), *The handbook of environmental psychology* (pp. 607–628). New York: John Wiley.

Kidder, L., & Fine, M. (1987). *Qualitative and quantitative methods: When stories converge* (New Directions for Program Evaluation, No. 35). San Francisco: Jossey–Bass.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Sadan, E., & Churchman, A. (1997). Process focused and product focused community planning: Two variations of empowering professional practice. *Community Development Journal*, *32*(1), 3–16.

Sandercock, L. (1998). *Towards cosmopolis: Planning for multicultural cities*. Chichester, UK: Wiley.

Sanoff, H. (2000). *Community participation methods in design and planning*. New York: John Wiley.

Sarkissian, W., Cook, A., & Walsh, K. (1997). *Community participation in practice: A practical guide*. Murdoch, Australia: Murdoch University, Institute for Science and Technology Policy.

Sarkissian, W., & Perlgut, D. (Eds.). (1988). *The community participation handbook: Resources for public involvement in the planning process* (2nd ed.). Murdoch, Australia: Murdoch University, Institute for Science and Technology Policy.

Schorr, L. B. (1997). *Common purpose: Strengthening families and neighborhoods to rebuild America*. New York: Anchor Books.

Wandersman, A., & Florin, P. (2000). Citizen participation and community organizations. In J. Rappaport, & E. Seidman (Eds.), *Handbook of community psychology* (pp. 247–272). New York: Kluwer Academic/Plenum.

Wates, N. (Ed.). (2000). *The community planning handbook*. London: Earthscan.

Environmental Psychology, Overview

Mirilia Bonnes and Giuseppe Carrus

University of Rome "La Sapienza," Rome, Italy

GLOSSARY

affordances Characteristics of environmental objects that are
related to both their physical properties and the possible
actions that each perceiver can perform with them.

conservation behaviors Actions that contribute to environ-
mental preservation and/or conservation.

environmental stressors Physical characteristics of the envi-
ronment that produce stress.

Hawthorne effect An increase in work productivity deriving
from a worker's awareness of being monitored by an external
observer.

participatory governance The transfer of authority and respon-
sibility from those who hold power by virtue of law, con-
tract, or organizational role to those not so empowered.

personal space The emotionally tinged zone around the
human body that people feel is "their" space.

place The product of physical properties of the environment,
people's cognitions of the environment, and people's
actions in the environment.

place attachment An affective bond between an individual
and a particular place that is not interchangeable with
another with the same functional quality.

privacy regulation The selective control of access to the self
or to one's group.

residential satisfaction The experience of pleasure or gratifi-
cation deriving from living in a specific place.

restoration The process of renewing, recovering, or re-
establishing physical, psychological, and social resources
or capabilities diminished in ongoing efforts to meet adap-
tive demands.

social dilemma A situation of intrinsic conflict between the
pursuing of individual gains and the maximization of col-
lective outcomes.

spatial cognition The internalized reflection and reconstruc-
tion of space in thought.

sustainable development Development that meets the needs
of present generations without compromising those of the
future.

territory A fixed geographical space marked and defended
by an organism and used for life-sustaining activities.

Environmental psychology is that branch of
psychology that studies the relationship between
people and the sociophysical features of the built and
natural environment, in order to enhance human well-
being and to improve people–environment relations. It
emerged as an autonomous field of scientific inquiry at
the end of the 1950s and during the 1960s. The main

conceptual roots of environmental psychology can be found in various theoretical proposals within psychology, all of them underlining the ecological or transactional basis and the context-specificity of human psychological processes. The emergence of environmental psychology is also related to the need for responding to specific demands from other technical and disciplinary fields, such as architecture and natural-ecological science, that face the problem of designing and managing the physical features of people's everyday environments. Due to the influence of these different factors, present-day environmental psychology can be characterized by four distinctive aspects: (1) the attention paid to the physical characteristics of the environment where human behavior occurs; (2) the wide variety of the research methods adopted; (3) the interest in problems with a clear social relevance; and (4) the interdisciplinary collaboration with other environmental fields. The first part of this article briefly outlines the basic historical and conceptual backgrounds of environmental psychology as a scientific discipline. Then, the main areas of interest and the related outcomes of current environmental psychology are briefly reviewed. (These aspects will be more thoroughly discussed within each of the single articles included in the present section.) Finally, perspectives on the future developments of environmental psychology are summarized.

environment, particularly when the built-up physical environment was considered. On the other hand, human behavior was conceived of as a "cause" of the physical environment, particularly when the natural environment was considered. However, this distinction was increasingly overcome by environmental psychology's growing emphasis on the necessity of adopting a transactional approach in the study of person–environment relationships. This approach lead to the person-in-place as the main unit of analysis for environmental psychological inquiry.

Various converging factors, originating inside, around, and outside the psychological field, contributed to the emergence and development of environmental psychology. The origin, the past, and the present of this branch of psychology can be better understood by looking at all these factors in order to outline a disciplinary identity that goes beyond the general label of applied psychology. Because of the influence of these different factors, present-day environmental psychology can be characterized by four distinctive aspects: the attention paid to the physical characteristics of the environment in which human behavior occurs; the wide variety of the research methods adopted; the specific interest in problems with a clear social relevance, and the adoption of interdisciplinary collaboration practices with other environmental fields devoted to the design and management of our everyday life environment.

1. THE HISTORICAL AND CONCEPTUAL BACKGROUNDS OF ENVIRONMENTAL PSYCHOLOGY

The starting date of environmental psychology as an autonomous disciplinary field within psychological science is commonly placed at the end of the 1950s. The environmental psychology that was formed during the 1950s and 1960s focused its attention on the physical features of the environment in which human behavior occurs. Its aim was to better understand the relationship between human behavior and the everyday physical or sociophysical environment. This environment was considered to be directly perceptible through the sensory organs, and was defined and considered in spatial and physical terms, whether built or natural, or on a small or large scale. At that time, there were two contrasting ideas about the relationship between human behavior and the physical environment. On the one hand, human behavior was conceived of as the "result" of the physical

1.1. Environmental Psychology within and around the Psychological Tradition

Psychology has been traditionally interested in environment–behavior interactions in a very general way. The major interest of environmental psychology rested on the discovery of the importance of the physical and spatial dimension of the environment as a constituting part of human actions and experience. This interest emerged within a wider ecological or transactional perspective coming from various domains of psychological inquiry.

Therefore, environmental psychology has always been concerned for the spatial-physical properties of the surroundings in which human behavior occurs, that is, of its physical setting. At the same time, it stressed the importance of considering these physical properties in a "molar" rather than "molecular" sense. It is also important to stress that the influence of the physical environment on human psychological processes often remains outside our individual and collective awareness.

Some pioneering studies in psychology highlighted the importance of the physical and spatial context in shaping human behavior, albeit often as an incidental outcome or a small part of other research aims. An example is the widely cited field experiments, conduced at the Western Electric Company by Elton Mayo in the 1930s, that investigated the effects of the lighting conditions in work settings upon workers' performance. Other pioneering works are the studies on the development of social influence networks by Leon Festinger and colleagues and the analysis of the "stream" of human behavior in natural settings by Roger Barker and colleagues. All of these earlier contributions were guided by a common methodological interest in studying human behavior in its natural setting, as researchers recognized the need for overcoming the usually low external validity of the traditional laboratory experiments. To this end, they preferred methods such as field experiments or non-obtrusive observation.

Other authors also played a crucial role in the emergence of environmental psychology, especially those who were more open to receive and develop ideas coming from areas that bordered on psychology, but that were traditionally interested in studying behavior in natural contexts. Areas such as cultural anthropology, about human proxemics, animal ethology, and microsociology are some examples. In other words, in order to be concerned with the spatial-physical environment, psychology had to get out from its habitual setting (i.e., the research laboratory), which was by definition a "non-environment." Just like as the earlier pioneers mentioned previously, these contributors also were generally opposed to the laboratory methods used by mainstream psychological research, and consequently were more willing to use other methodologies such as field experiments and field observations, both natural or systematic. R. Sommer and I. Altman's studies on personal space and social behavior, which remain as cornerstones of the early environmental psychology, are good examples.

The enthusiasm over the emergence of this new field of inquiry was then linked with the previously mentioned dissatisfaction for both the poor ecological validity and the low social relevance of much laboratory research. This resulted in a search for a "real world psychology." This frequent dissatisfaction can be traced to the various forms of ecological demand raised since the 1940s and 1950s by various authors and psychological schools, which later developed into what has been called "contextualism" or "contextual revolution." This revolution is certainly at the core of the development of environmental psychology, particularly in its transactional-contextual approach, which has characterized environmental psychology since its beginning.

Initially, the increased awareness of the crucial effect played by the physical features of the everyday environment on human behavior was based on two main theoretical psychological traditions. The first is the psychology of perception, especially in its more ecological orientations, such as the lens model by E. Brunswik, the transactional school of the Princeton group, and the ecological approach to perception of J. Gibson. Gibson, for example, introduced the neologism affordances, which identifies the physical properties of environmental objects that are related to the possible actions that can be performed with them. The second psychological tradition is based on the social psychology approach, through the pioneering work of K. Lewin, E. Tolman, R. Barker, and U. Bronfenbrenner. It embraces a more "holistic" or "molar" perspective, which later developed into the transactional-contextual approach to person–environment relationships, as systematically outlined by many contributors to the first *Handbook of Environmental Psychology* edited in 1987 by D. Stokols and I. Altman. In this perspective, the physical environment or physical setting has been increasingly considered as a sociophysical environment, with a growing emphasis on the social aspects of both the physical environment and the psychological processes involved. The notion of place, and its related environmental-psychological processes, also became a central unit of analysis for many environmental psychologists. Typically, places were defined from an environmental psychology point of view as a product of three main dimensions: physical properties, people's cognitions, and people's actions.

1.2. Environmental Psychology and Other Environmental Fields

The problem-oriented demands rising in technical and disciplinary fields distant from psychology are also an important factor that contributed to the emerging of environmental psychology. Examples of these fields are architecture, engineering, urban planning, human geography, natural and ecological science. Typically, architects and engineers are concerned with problems regarding the relationships between people and the built or "human-made" environments. In contrast, geographers and natural scientists are more interested in the relationships between people and the natural features of the environment.

In architecture and urban planning, those who were dissatisfied with an egocentric approach to design desired to move toward a user-centered approach to design, as well as to move from "product" design and planning to "processes" design and planning. A seminal work in this area is that of the urban planner Kevin Lynch at the Massachusetts Institute of Technology. His famous book *The Image of the City,* published in 1960, is considered a cornerstone of the collaboration between urban planning and environmental psychology. Lynch proposed that the point of view or the image that users form of the urban environment should be kept in mind when designing and planning urban spaces. Likewise, in engineering and technology, many scholars and practitioners became interested in the human-use dimension of technological systems.

The collaboration between psychology and architecture was mainly guided by an interest in the influence of specific and localized spatial characteristics of the built-up environment on human behavior. A different orientation, however, characterizes the interest in human behavior that arose in the fields of natural and ecological science, which is the other main external domain contributing to the establishment of environmental psychology. Here, the focus was placed on the possible (and usually negative) impact that human behavior can have upon the natural environment, at both a local and a global or biospheric level. Therefore the importance of paying attention to the human dimension—considered at an individual, social and cultural level—of global environmental changes, such as the greenhouse effect and climate changes, the loss of biodiversity, the depletion of the ozone layer, and so forth, was recognized. In parallel with the increasing relevance assumed by these environmental problems, ecologists also became increasingly aware of the need to integrate the natural and social sciences when dealing with environmental phenomena. The growing interests of current environmental psychology in topics such as environmental concern, pro-environmental values and attitudes, ecologically relevant behaviors, sustainable lifestyles, and natural resource management is also a consequence of this increased awareness.

2. THE MAIN AREAS OF INTEREST OF CURRENT ENVIRONMENTAL PSYCHOLOGY

Contributions in environmental psychology can be grouped into five general domains: spatial behavior,

environmental cognition, environmental stress and restoration, environmental assessment, and environmental concern and resource management. In this section of the article, the main empirical results and outcomes of environmental psychology are briefly introduced for each of these five domains. The various single articles composing this section provide a more exhaustive treatment of the different topics related to these five domains.

The environmental psychological processes grouped into these five domains can be considered transverse to the variety of specific settings in everyday life. However, since environmental psychology has also been increasingly characterized by the adoption of a place-specific approach to people–environment interactions, many environmental psychologists have been increasingly interested in a number of specific places that have a particular impact on everyday life and well-being, such as offices, schools, houses, stores, hospitals, museums, prisons, cities, etc. Three articles in this section focus on three particularly important places in our everyday experience: the workplace, the school, and the city.

2.1. Spatial Behavior

The concept of spatial behavior relates to how individuals regulate and use (in terms of appropriation and defense) their spatial environments at different personal, interpersonal, and group levels. Environmental psychology interested in spatial behavior focuses on the role of the spatial properties of the environment in shaping and regulating social interaction in everyday situations. Four major environmental psychological concepts are considered in this field: territoriality, personal space, privacy, and crowding. The articles by R. Sommer and by C. Werner, B. Brown, and I. Altman review the relevant empirical findings of these domains of environmental psychology, as well as their implications for daily life.

These domains encompass the variety of strategies by which people and communities set up and regulate the spatial boundaries of their living environment (territoriality) and the very personal sphere of the amount of space we put between us and the others (personal space). The concept of privacy encompasses both of these aspects, as it refers more generally to the various personal and group-based processes through which people set and control their mutual closeness in daily interactions.

The article by R. Sommer thoroughly illustrates the concept of human territoriality. As Sommer outlines, territoriality is the spatial appropriation, marking, and defense of our living spaces. Following an evolutionary orientation, territories can be defined as those spatial

areas that deserve to be defended from external intruders. An intriguing issue connected with human territoriality is that of crime prevention, through the concept of defensible space. As the article describes, there is a widely documented negative correlation between the presence of physical and symbolic territorial markers (e.g., well-maintained dwellings, tended yards, signs of occupancy) and the likelihood of criminal intrusion and vandalism in residential areas. The relationship between the presence of markers and crime, however, may be mediated by social variables. Notably, clear territorial markers are related to stronger residential identification, place attachment, sense of community, and higher feelings of safety among residents. Thus, architectural defensive features might play an indirect role in reducing the likelihood of crime. They reflect and communicate strengthened community-based ties and greater social cohesion; these make potential intruders less willing to perpetrate criminal activities. Another important outcome of the concept of territoriality is related to urban planning. Sommer's article in fact points out how specific design features encouraging residents' identification with and control over a specific territory may promote its proper maintenance. Furthermore, common areas such as urban parks or playgrounds should not be placed between the territories of rival gangs, as they would then run the risk of being abandoned and vandalized. Other domains in which the study of human territoriality has offered useful insights are sport (as sport teams perform better on their home field), police investigations (as many criminal gangs mark their territories by the use of graffiti and other symbols), and environmental conflict resolution (as the creation of ad hoc territories reserved to different stakeholders may prevent the emergence of land use disputes).

A second article by R. Sommer illustrates the environmental psychological concept of personal space. As opposed to territoriality, which is a place-based or site-based concept, personal space is a person-related, and thus trans-place, concept. In his seminal studies in this field, Sommer defined personal space as that emotionally tinged area that people desire to maintain around themselves and that they feel is "their space." Sommer's article also provides a review of the various individual, situational, group, and cultural factors regulating personal space. As this article reveals, personal space has relevant implications for several domains of human–environment interaction. For example, the personal space concept has provided insights for the design of institutional settings with fixed seating, as well as in the domain of mass transportation. Likewise, issues

regarding personal space were incorporated in the design of public or semi-public settings (e.g., offices, stores, banks). These should be set up in order to provide spaces that leave the users as free as possible to shift among different degrees of desired mutual closeness. Many guidebooks describing the appropriate spacing in different social encounters are also available. For example, salesman are trained about how their selling activities might profit from an appropriate spatial interaction with customers. Finally, the issue of personal space has serious implications for the legal field: for example, the invasion of personal space has been an issue in court cases concerning prison crowding and sexual harassment. It is also an aspect considered during jury selection in high-profile trials, where professional consultants advise lawyers in order to detect potentially biased jurors.

Personal space and territoriality are both involved in the more general concept of privacy regulation. In this sense, they can be seen as two different strategies that people use to regulate their privacy, and through which people strive for psychologically satisfactory level of openness or closedness to others. A seminal contribution to the definition of privacy in environmental psychology comes from I. Altman. Within this section, the reader will find a detailed presentation about the psychological mechanisms of privacy regulation in the article written by C. Werner, B. Brown, and I. Altman. As they report, privacy regulation is defined as the process of selectively controlling access to one's self or group. In their article, Werner et al. underline how environmental psychology treats privacy as a dialectic process: that is, people might avoid or seek social contact, depending on the specific situation they are in. An important point is that individuals feel more comfortable when allowed to be as open or as closed to the others as they desire. This is also related to another important issue in the environmental psychology of spatial behavior, that is, crowding. As reported by Werner et al. in their article, crowding can be defined as one's perception that there are too many people present in a given situation. The possibility of a proper privacy regulation therefore becomes very important when people experience crowding. Environmental psychology has also treated crowding as a specific environmental stressor.

The article of Werner, Brown, and Altman also refers to how environmental psychology has identified a number of physical features that either allow people to experience a satisfactory regulation of privacy or do not. As a consequence, the privacy concept has many practical implications for the design of various specific environments such as workplaces, schools, prisons,

hospitals, and public residences. For example, university dormitories designed as suites or apartments can promote the formation of small subgroups of occupants; this is meant to buffer the possible negative consequences of crowding upon the possibility of privacy regulation. When designing the layout of a home, architects should place rooms that are usually considered more private (e.g., the bedrooms, or the family bathrooms) far from the entrance. Likewise, privacy-related concepts have been applied to urban planning and community development. The article illustrates how, for example, public spaces provided with design features encouraging interpersonal contacts (e.g., streets that invite walking, tree-shaded streets, low speed limits, green areas) might be used to promote strengthened community bonds and social ties among neighbors.

Beyond the effect that specific design features can have in affording or impeding people's mutual interactions, a central factor in people's satisfaction or dissatisfaction with privacy conditions is the possibility of regulation and control. A general recommendation is that the physical environment should be designed in order to provide to its users the possibility of regulating privacy, by seeking or avoiding social contact according to their situational and personal needs and desires.

2.2. Environmental Cognition and Cognitive Mapping

Since its beginning, environmental psychology has focused on the relationship between cognitive processes and physical space. The article of R. Golledge on spatial cognition thoroughly reviews this area of study. As Golledge notes, spatial cognition is a multidisciplinary area involving psychological science (e.g., cognitive and developmental psychology, neuropsychology) as well as other disciplinary and technical fields (e.g., geography, anthropology, architecture and urban planning, computer science).

A core concept in the environmental psychology of spatial cognition has been cognitive mapping. The term cognitive mapping refers to the process of mentally acquiring, representing, storing, and using information about the spatial properties of the physical surroundings and the relations among its constitutive elements. The interest in cognitive maps in psychology can be traced back to the works of E. Tolman in 1948. Studies on cognitive mapping have focused on the different kinds of information people rely on when forming a map of a given spatial setting.

Studies on cognitive mapping have dealt with the representation of large-scale urban environments, as well as with the representation of the interior of buildings. One goal of this kind of research was to investigate how to aid people's orientation, memory of places, and wayfinding in complex environments.

Kevin Lynch's seminal book *The Image of the City* influenced several ensuing works on the formation of mental maps of urban settings. Urban planners can use principles highlighted by these studies in order to design urban spaces that are capable of facilitating people's orientation in the city. For example, some spatial features of the residential environment, such as spatial prominence, sharp contours of buildings, easy accessibility, and the presence of natural surroundings might facilitate the overall comprehension and use of neighborhoods. Likewise, principles drawn from the cognitive mapping literature have been used for aiding wayfinding in complex interior structures, such as university buildings, hospitals, or museums. The presence of proper signs (floor plans, large graphic aids) and the availability of visual accesses may be effective in enhancing users' proper orientation.

As Golledge points out in his article, there is now a particular interest in the study of spatial cognition in aged people. The increased trend of aging in western societies has in fact led to an increased concern for the quality of life of older people. Research in this field is therefore being devoted to trying to identify how the layout of the physical space can be designed and managed in order to buffer the declines in spatial competence associated with aging, and the negative consequences of this decline for the safety, effectiveness, and autonomy of the elderly.

2.3. Environmental Stress and Restorative Environments

The environment in which most of humankind currently lives is often characterized by the presence of several potentially adverse physical conditions that can be a chronic, powerful, and uncontrollable source of psychological distress. The article by G. Evans and S. Cohen on environmental stress, included in this section, provides an exhaustive review of this traditional area of study of environmental psychology.

Environmental stressors can be defined as those actual or perceived adverse properties of the physical environment that are capable of producing a negative (physiologically and psychologically costly) effect

upon a person. It is important to note that one's everyday environment is often a source of simultaneous and interdependent multiple stressors. The article by Evans and Cohen addresses this particular aspect of environmental stress. The effects of different environmental stressors frequently add or interact with each other. As Evans and Cohen outline, individuals must make a remarkable adaptive effort in order to cope with such adverse conditions. Because humans have great powers of adaptation in adopting strategies to cope with stressors, the coping activity may become a stressful condition itself, particularly in the long-term. The article focuses on the four main costs of adapting to poor environmental conditions: cumulative fatigue, learned helplessness, physiological mobilization, and overgeneralization. These adaptive costs are caused by chronic exposure to a number of different environmental stressors, such as crowding, noise, or pollution.

Some of these main environmental stressors can be a very common experience for a great number of people, in particular for residents of cities. Conversely, there are other critical environmental conditions that are very uncommon that are also of interest to environmental psychologists. The article by R. Bechtel on extreme environments and mental function elaborates on this issue. As Bechtel points out, there are two principal classifications of extreme environments. On the one hand, there are permanent extreme environments, such as high mountains, deserts, cold regions, and jungles, and on the other hand, there are temporary extreme environments, such as Antarctic polar stations, space shuttles, and simulation laboratories. The main goal of environmental psychology concerned with extreme environments is to better understand how human mental function and behavior can be affected by these kinds of environmental conditions in order to buffer their potentially negative impact on mental function: that is, how to make these environments less extreme for their users.

In recent decades, environmental psychology has also addressed stress-related issues from an opposite viewpoint: that is, how the physical environment can restore human mental functions, and therefore promote psychological well-being. The article by T. Hartig on restorative environments provides an exhaustive review of this promising literature. As the article explains, restorative environments are those that not only permit, but promote, restoration, where restoration is defined as the process of recovering physical, psychological, and social resources being diminished by efforts to adapt to external demands.

In his article, Hartig reports how a considerable amount of empirical research has demonstrated the existence of specific environmental conditions capable of positively affecting people's feelings of well-being, although different theoretical elaborations have been proposed to explain it. Perhaps the most well-documented effect within this area of research is that natural settings appear to be more restorative than built-up ones, because they stimulate people's interest and provide people with the possibility of being away from their usual experience. Furthermore, natural environments are highly restorative because of various specific features, such as visual depth and visual complexity.

Research on restorative environments can have a high relevance for the design of health care structures. A widely cited study by R. Ulrich, published in *Science* in 1984, showed for example that providing hospital patients with the possibility of contact with nature (even just visual contact) might positively impact the effects of medical therapies. Restorativeness is a relevant issue for the domain of urban planning as well. Setting up urban green areas that are easily accessible to urban dwellers can provide them with more frequent opportunities for psychological restoration in a highly stressful environment such as the city.

2.4. Environmental Assessment

Although a theoretical and conceptual distinction between environmental assessment and environmental appraisal has been proposed, the former being more place-centered, the latter being more person-centered, here both of these aspects of assessment are considered. Environmental psychology interested in environmental assessment has encompassed a large amount of empirical contributions and theoretical models aimed at answering very important questions. For example, how do people evaluate, judge, or express preference for different kinds of environments? How do the physical features of a setting interact with people's personal and social characteristics when they evaluate an environment? In particular, which physical properties of the environment are related to positive evaluations and which are related to negative ones? Furthermore, what are the individual and group differences that may lead to differing environmental evaluations? And also, how do cognitive and affective judgments interact when a certain environmental scene is evaluated?

Generally speaking, different affective qualities have been found to characterize different environments. These qualities vary according to two main bipolar axes, pleasure (pleasant/unpleasant) and arousal (high/low). The physical surroundings of people's everyday

lives can generate positive or negative emotions in terms of the pleasure they afford to the perceiver, or in terms of the amount of arousal they provide to the perceiver.

According to a more cognitive model proposed by R. and S. Kaplan, there are four main physical characteristics predicting the pleasantness or unpleasantness of an environmental setting: coherence, complexity, legibility, and mystery. Coherence and complexity are two features of environmental scenes that immediately strike the perceiver. Conversely, legibility and mystery are two features that can be inferred after viewing an environmental scene for more time. The Kaplans' model also offered interesting insights into the study of landscape preference.

Whether a setting will generate positive or negative emotions in its users/perceivers is also a function of the prior state of the perceiver him- or herself. People's affective appraisal of a given environment is an adaptive process, in the sense that it is a function of both the properties of the environment and the characteristics of the perceiver. For example, high sensation seekers prefer highly arousing environmental scenes, and vice versa.

Several studies have also showed that natural environments are usually preferred to built environments by a large number of subjects. This preference seems to be consistent across different cultures, genders, ages, and so forth. Regardless if they are evaluated in terms of general preference, aesthetic beauty, or perceived restorativeness, natural landscapes are judged more positively than built settings. The presence of water in an environment is also positively judged. Moreover, as previously mentioned, exposure to natural views has been proved to have positive effects upon physiological and cognitive indicators, such as the rapidity of healing after hospitalization and the relief of mental fatigue. Such a nature favoritism, or biophilia, has been explained in evolutionary terms: people are still attracted by nature because it provided sources of sustenance and shelter in the earlier stages of human evolution. The psychological benefits of contact with nature can be due to the fact that evolution favored those who could benefit more from exposure to nature.

Despite this apparently generalized nature favoritism, a limited degree of human intervention within a natural environment is associated with positive evaluations. That is, scenes in which the human presence is visible (although not too intrusive) are usually preferred to scenes of total wilderness. This leads to the importance of another factor in shaping environmental evaluations, perceived control. People tend to judge more positively those environments that they perceive as controllable, compared to those that they perceive as uncontrollable.

Typically, the perception of control is associated with higher perceived safety. Speaking more generally, environmental evaluation could be conceived of as a function of the perceived cognitive, affective, and behavioral fit between the characteristics of the person (e.g., needs, goals, values, expectations) and the actual physical properties of the environment. In other words, people will prefer those environments they see as matching their needs, goals, and expectancies in that specific context at that given time.

The processes driving environmental assessments are relevant for the study of residential preference and satisfaction. The article written by M.V. Giuliani on residential preference and attachment across the lifespan, and the article by M. Bonaiuto on residential satisfaction and perceived residential environment quality, both included in this section, specifically deal with these issues. As both Giuliani and Bonaiuto describe in their articles, residential preference and satisfaction are multidimensional processes, given the tendency of people to combine different aspects when judging the quality of a residential setting. At a macro-level (e.g., the neighborhood) these aspects are related to (1) the physical and aesthetic characteristics of residential areas, (2) the social relations that people can establish in a specific residential area, and (3) the actions and behaviors that people can perform in a specific residential area. At a micro-level (e.g., the home) these aspects are related to (1) the quality of the interior arrangement and decoration of the house, (2) the specific architectural style of the house, and (3) the location of the house, in terms of both proximity to functional services and distance form the city center.

An important outcome of studies on environmental assessment has been the development of psychometric tools for the measurement of environmental preference and satisfaction, such as perceived environmental quality indicators (PEQIs). Because environmental quality assessment is conceived as a multi-dimensional process, PEQIs should encompass a sufficient variety of dimensions: the most widely shared qualities are aesthetics, functionality, safety, social relations, noise and pollution, and green spaces.

The information gathered through these tools can often represent the first step in developing more inclusive and participatory methods in environmental design and planning. Why is including users' perspectives or letting people take part in decisions considered so important in the environmental domain, as well as in many other domains? The article by A. Churchman and E. Sadan on participation in environmental design and planning thoroughly illustrates this issue. As Churchman and

Sadan outline, the major argument in favor of public participation is that it enables decisions that better fit the needs of the people. The positive consequences of this better fit are twofold, as both public support for environmental policies and people's care for a proper maintenance of their environments will be more likely. Nonetheless, Churchman and Sadan also identify and discuss a number of factors that may bedevil the effective implementation of participatory processes in the environmental domain. For example, public leaders and politicians sometimes see participation as a threat to their power. It is important to outline here how this article presents various arguments that support the crucial role of environmental psychology in promoting and improving participatory processes in environmental design, planning, and management. This aspect is particularly relevant, as it relates to the issue of improving the processes of inclusive environmental governance, which is currently seen as a crucial goal by many international and intergovernmental authorities.

2.5. Environmental Concern, Environmentally Friendly Behaviors, and Natural Resources

A shared belief among the scientific community and among public opinion is that the quality of our environment at the local and global level has rapidly decreased almost everywhere in the last decades. As a matter of fact, the major reason for that progressive and often dramatic deterioration is the impact of many human activities and rapid industrial growth on the ecosystems. Therefore, the healthy or unhealthy state of the environment primarily depends, and will depend in the future, on those human activities that can cause relevant and often irreversible environmental local and global changes.

In order to deal with these kind of problems, environmental psychology has increasingly addressed the issue of environmental concern in the last two decades. Two articles included in this section illustrate this important area of interest in present-day environmental psychology.

In his analysis of conservation behavior, F. Kaiser focuses on the psychological factors at the basis of the actions contributing to environmental preservation. As Kaiser suggests, people's conservation behaviors can be approached from two rather different perspectives. Conservation behavior, in fact, can be defined either from an observer's or from an actor's viewpoint. The first implies a focus on the consequences for the environment and considers apparently similar actions (e.g.,

recycling paper and recycling used batteries) as distinct behaviors. The second perspective focuses more directly on the human motivations driving conservation behaviors and groups even apparently different actions (e.g., recycling glass and owning solar panels) into the same behavioral category. Kaiser's article also illustrates how this second approach is more directly linked to a psychological perspective. On the contrary, a too strict focus on environmental consequences may lead to underestimating the psychological determinants of conservation actions.

But what are the psychological determinants of environmentally friendly behaviors? The article by H. Staats on pro-environmental attitudes and behavioral change deals more directly with this question. In particular, this article focuses on the relationship between environmentally friendly attitudes and behaviors: that is, the congruence or incongruence between what people believe and feel and how they behave toward the environment.

Pro-environmental attitudes are increasingly shared by many people in western societies. Some social structural variables (e.g., age, gender, level of education, place of residence, political orientation) appear to be related with this concern. Typically, young, female, highly educated, urban, and liberally oriented subjects express a stronger concern for the environment. The increased concern for environmental issues has been summed up by the concept, introduced by R. Dunlap and K. Van Liere at the end of the 1970s, of a new environmental paradigm (NEP). The core idea of NEP is that an increased number of people are developing a new perspective toward the environment, based on the belief that the state of the earth's ecosystems is becoming more and more precarious, therefore compromising human survival.

However, people often do not coherently translate such a positive concern into consequent pro-environmental behaviors. The reason for this lack of correspondence has been explained in different ways. Staats' article guides the reader through two important social psychological models of attitude-behavior relations: the theory of reasoned action (TRA) and the theory of planned behavior (TPB). As Staats points out, the basic tenet of these models is that the relation between attitudes and behaviors is mediated or moderated by other variables, such as behavioral intentions and perceived behavioral control. TRA and TPB have been applied, sometimes successfully and other times less successfully, to several environmentally relevant behavioral domains, such as waste recycling, travel modes, water use, energy use, and green consumerism.

At any rate, the role of variables other than attitudes seems to be crucial for explaining pro-environmental actions: among them there are past behaviors and habits, environmental knowledge, social and personal norms, and value orientations. Well-established habits are often difficult to quit, even when their performers are aware of the possible negative consequences for the environment. Likewise, a lack of specific knowledge about environmental issues, or about the environmental consequences of a specific behavior, may function as a barrier to pro-environmental actions: people are in fact frequently unaware of, or uncertain about, the actual state of the environment and the negative consequences of their behaviors on it. For example, a deeper knowledge of environmental issues seems to distinguish committed environmental activists from the general public, although it is difficult to state whether knowledge comes before activism, or vice versa. Both social and personal norms are also related to pro-environmental behaviors. Although some explicit pro-environmental norms are socially shared, people might not always behave according to these social norms, especially because several, sometimes conflicting, norms can be present at the same time in a given situation. In particular, the divergence between prescriptive (what is explicitly prescribed) and descriptive (what is observed in others' behaviors) social norms can hamper pro-environmental actions. A series of interesting experiments conduced by R. Cialdini and his colleagues showed that the context in which prescriptive and descriptive norms are framed and made salient can account for a particular environmental behavior such as littering. The role of values and ethical principles in shaping pro-environmental behaviors has been explored as well. A distinction has been proposed between two major value orientations about environmental issues: ecocentrism and anthropocentrism. Ecocentric people value environmental preservation for its own sake; conversely, anthropocentric people value environmental preservation because of the positive consequence that it can have upon human well-being. The former are more likely to behave coherently with their pro-environmental attitudes compared to the latter, even when this implies some personal or economic cost.

But what are better methods for the promotion of environmentally responsible behaviors and for the prevention of environmentally adverse behaviors? In his article, Staats outlines various possible strategies that seem to be effective to this end; they can be information-based (e.g., communication campaigns, education,

advertising), incentive-based (e.g., monetary rewards/punishments), prescription-based (e.g., laws, rules, regulations), or community-based (e.g., public involvement and participatory programs). Different kinds of behaviors in different contexts deserve different kinds of strategies. As P. Stern has recently pointed out, there are some basic principles to be followed for enhancing the likelihood of changing environmentally relevant behaviors: the use of mixed intervention strategies, the adoption of an actor's perspective, the constant monitoring of programs, and the promotion of public participation.

The models trying to explain the lack of correspondence between pro-environmental attitudes and behaviors that were briefly reviewed above rely on the role of various factors that could interfere with the supposed positive association between attitudes and behaviors. This theoretical position, however, may not adequately account for other social and societal processes involved in people–environment relations, in particular, the mutual interdependency of people's actions in the environment. A theoretical approach that delves more directly into this direction is the social dilemma paradigm.

Social dilemmas (SDs) are situations in which two or more people are faced with a choice between pursuing individual gains or maximizing collective benefits: if everyone (or too many) chooses the former option, everyone is worse off than they would be if everyone (or sufficiently many) chose the latter. In other words, in a SD, the interest of a single person is in conflict with a more general collective interest, and one is faced with a choice between a defective (or selfish) option and a cooperative (or altruistic) alternative. Such a conflicting situation often exists in the environmental field when individual and collective interests clash over the exploitation of a limited natural resource. These dilemmas are usually referred to as commons dilemmas.

As Staats suggest in his article, almost every environmental problem can be framed and approached as a commons dilemma situation. Most of the times the individual-collective conflict in the exploitation of a limited resource has a temporal dimension as well. That is, maximizing individual gains in the short term will result in a collective loss in the long run (the common resource will be extinguished); conversely, limiting the individual gains in the short term will result in a collective benefit in the long run (the common resource will be guaranteed). This process is at the core of the political concept of sustainable development, and has been well illustrated by G. Hardin in his famous 1968 article entitled "The Tragedy of the Commons."

On the basis of the SD paradigm, it is not surprising that people frequently find out that behaving in an environmentally negative way is in the short term more advantageous, easier, more comfortable, less costly, and so forth. To an extreme extent, in many daily situations, environmentally unfriendly behaviors are the rule and not the exception.

In addition to the Staats' article, two other contributions included in this section deal, more or less directly, with the SD paradigm. In particular, in his article on environmental versus individual risk taking, C. Vlek reviews the literature on individual risk taking in order to understand how to manage environmental risks as a collective risk. As Vlek illustrates, various traditional psychological theories on risk can be considered when dealing with environmental risk perception, such as classical decision theory, prospect theory, risk omeosthasis, protection motivation theory, and emotional approaches to risk. Furthermore, it is shown that many collective risks for the environment are often the result of the summed external negative effects of numerous individually advantageous activities. To this extent, the SD paradigm is directly called into question by Vleks' article, as it shows how the minimization of environmental collective risks and the maximization of individual benefits may be incompatible goals in many everyday situations.

Likewise, the article on travel and the environment by T. Garling illustrates how concepts deriving from the SD paradigm can be relevant to the domain of motorized transportation. As Garling shows, the increasing trend of private car use in urban areas is starting to pose serious threats to the global environment and to the health and well-being of urban dwellers. These threats are overwhelming the individual benefits of the single car users. The choice of whether to shift from private to public motorized transportation systems or from motorized to un-motorized vehicles can be framed as a social dilemma for many car and motorbike drivers. Therefore, a number of policy measures aiming at reducing private car use among urban dwellers that are outlined by the Garling's article can be seen as attempts to solve that dilemma.

But which are the main factors orienting our choices in social dilemma situations? Three crucial components of SDs are information, communication, and social identification. The possibility of monitoring others' behaviors and the possibility of communication may promote cooperation. Likewise, high social identification may result in stronger cooperation, although the opposite may be true. In fact, when a dilemma involves two or more groups competing for the same resource,

identification with a specific group may result in diminished cooperation.

These aspects may also have relevant implications for legitimizing and facilitating the acceptance of authority structures regulating citizens' environmental behaviors in our societies. The way a situation is structured by the authorities and interpreted by the individuals acting in it is crucial for approaching environmental dilemmas. An example can be found in the domain of biodiversity conservation. Here, highly centralized environmental policies (e.g., land use regulations, designations of National Parks) can represent a dilemma in the perception of local stakeholders. A strong sense of local identification might then form a basis for in-group biased environmental perceptions, in which public authorities are seen as a conflicting out-group by local identifiers. This can lead to the emergence of local protests against specific environmental regulations.

In sum, the SD paradigm is increasingly being used to address many current environmental problems regarding the management of limited natural resources. The results from studies based on this paradigm suggest that strategies for solving collective environmental dilemmas in our societies should stress the common fate between individual group members, in particular because they show the crucial importance of communication and shared identities among individuals for the promotion of cooperative and pro-environmental choices.

2.6. Place-Specific Environmental Psychology

The articles in this section described above refer to environmental psychological processes that can be considered as transverse to the places where people's everyday experiences occur. However, current environmental psychology has been increasingly characterized by the adoption of a place-specific approach to people–environment interactions. This has lead many environmental psychologists to concentrate their interest on specific places, ranging from the micro-scale of home interiors, to the meso-scale of workplaces, schools, and hospitals, to the macro-scale of neighborhoods and cities. There are three articles in this section that focus more directly on three specific places considered as particularly relevant to many people's daily lives: the workplace, schools, and the urban environment.

The article written by R. Walden on working environments reviews some of the main theoretical and empirical contributions addressing environment–behavior

relations in the workplace. In her contribution, Walden notes how the study of the physical characteristics of workplaces can be traced back to the beginnings of psychological inquiry. The studies by E. Mayo and his coworkers at the Western Electric Company on the effects of the physical environment upon work performance date back to the first half of the 20th century. These pioneering works clearly highlighted how the physical features of the work settings can interact with the interpersonal relations and organizational aspects at work in shaping work performance. Present-day models of environment–behavior relations in the workplace, such as one recently proposed by R. Gifford, still reflect the influence of Mayo's earlier findings. The model by Gifford mentioned in Walden's article places a particular emphasis on the different degrees of congruence or incongruence resulting from the interaction between the personal characteristics of workers (e.g., experience, job level, personality, skills, motivation) and the psychical characteristics of the work setting (e.g., noise, temperature, light, spatial density, quality of materials). The higher the congruence, the better the outcomes of the working activities. These can be assessed in terms of workers' stress or well-being, individual and collective performance, job satisfaction, or organizational commitment. In addition, various psychological processes (e.g., arousal, adaptation, overload, emotions, personal control) may act as mediators of the relationship between the environment–person fit and the outcomes of the working activities. As Walden notes in her article, future trends in the design of workplaces point to rapid development of communication technologies as well as to the role of these in matching the augmented concern for sustainability that characterizes human societies. One example is the possibility for off-site working, or telecommuting, through the use of remote cable connections or wireless systems. This possibility could partly offset the disadvantages of traffic congestions among commuters.

A second article by R. Walden, on school environments, focuses on another place that has also been widely investigated by environmental psychology. As with workplaces, the attention devoted to schools is a long-standing tradition in environmental psychology. In fact, the interest of environmental psychology for schools can be traced back to the pioneering studies of Barker and colleagues in the 1960s. Generally speaking, environmental psychology has clearly outlined how the physical features of schools have a considerable impact on the effectiveness of the educational

program. Different physical characteristics of schools can either promote or hamper the education process. As Walden reports, there is consistent evidence showing that small schools do a better job than large schools. Empirical research has showed how smaller schools are characterized by a better relationship between students and teachers, diminished rates of vandalism and violence, faster career advancement of children with disadvantaged backgrounds, increased participation in and commitment to school activities, and greater satisfaction of parents. Therefore, Walden states that schools with no more than 500–700 pupils appear to be a reasonable solution, while schools with more then 2100 students have shown the worst performances. Apart from school size, several other characteristics of the educational physical setting may affect the quality of the education process. The most important are the state of maintenance of the school buildings, the presence of design features affording proper privacy regulation, good acoustics in classrooms, the availability of daylight in classrooms, the presence of playgrounds and schoolyards furnished with nature and natural materials, and the presence of carpets and decorations. Finally, the design of new schools can benefit from the involvement of students, teachers, parents, and members of the local communities.

A third article focuses on another important location that encompasses the two previously mentioned places: the city. As G. Moser outlines in his contribution on urban environment and human behavior, about half of the world's population spends the majority of their lifetime in big cities.

As reported by Moser, cities are more and more becoming places in which daily conditions are far from optimal for their inhabitants: cities are by and large considered as bad places to live. In line with this trend, environmental psychology has introduced the concept of urban stress to indicate the sources of stress that are typical of urban life, and therefore that affect a large percentage of the population. Living in a large-scale urban settlement can frequently become a source of multiple stressors, because of the simultaneous presence of several adverse environmental conditions. This aspect has also been discussed in the Evans and Cohen article on stress. For example, urban settings can amplify the relationships among crowding, social withdrawal, social support, and pro-social behaviors. Moreover, urban life presents several conditions in which people perceive a very low degree of control over events: this often results in diminished self-efficacy and helplessness.

On the other hand, despite their bad reputation, cities continue to attract a large part of the human population. Therefore, as the Moser's article reminds, it is not surprising that urban settlements have been a specific object of analyses for many scholars in environmental psychology. Some important psychological processes related to the everyday urban experience were actually pointed out by environmental psychology, such as place- and urban-related identity and urban attachment. In that sense, cities are a primary arena in which almost any environmental psychological theory and empirical finding can be considered. As a consequence, cities can be thought of as a wide field-laboratory for testing environmental psychological theories. But what are the main trends of environmental psychological research on urban settlements? The article by Moser describes the increased concern for issues of sustainability and sustainable development that have emerged in the scientific community in the last decade. It is worth stressing how cities, considered as ecosystems, were seen as having a crucial role by the many intergovernmental agencies concerned with the promotion of sustainability. As a consequence, the concept of sustainable cities has been proposed in order to define the major environmental transformations needed to improve urban systems in order to enhance their ecological "performance." These issues are the next major challenge for environmental psychology focused on urban issues. Embracing this perspective, Moser also shows how urban-related identity is an important environmental psychological factor that may account for the adoption of sustainable lifestyles among city dwellers. As an example, Moser mentions the city-identity-sustainability (CIS) model that has been recently proposed by E. Pol and others. According to the CIS model, a strong sense of identification with the city can lead urban inhabitants' to a stronger and more durable engagement in pro-environmental activities.

3. THE FUTURE OF ENVIRONMENTAL PSYCHOLOGY

Since its beginnings, environmental psychology has been a frontier field within psychological science. On the one hand, following the ecological demand coming from several fields of psychology, environmental psychology has tried to underlie the importance of a contextual/transactional approach to psychological inquiry. Within this approach, environmental psychology has focused on an aspect that was typically underestimated by mainstream psychology: the role of physical features of everyday life environments in affecting human psychological processes, behavior, and well-being. On the other hand, environmental psychology has tried to respond to specific and urgent demands coming from various technical fields devoted to the design, change, preservation, and management of these everyday environments.

In many aspects, environmental psychology is still a frontier field of psychological inquiry, as its primary aims have not substantially changed over the decades. The whole story about the context-related character of human behavior and mental function is still far from being known due to the wide variability and diversity of the physical contexts in which human behavior takes place. At the same time, many of the urgent environmental problems that generated an external demand for the development of environmental psychology have not been solved yet. Some of them have become even more serious, as the quality of the environment and the availability of life-supporting resources have diminished consistently in many parts of the world.

Thus, the research agenda of environmental psychology has broadened considerably in the last 20 years. For example, the collaboration between psychologists and architects has increasingly included people–natural environment relations together with people–built environment relations. The domain of political ecology with its concept of sustainable development has influenced the fields of architecture, engineering, and urban planning. The increased relevance of various new technological and architectural domains such as bioarchitecture and landscape architecture is an example of such an influence. Also, an increasing focus has been placed on the human dimension of global environmental and biospheric changes.

To address these issues, many different psychological theories can be used, coming from a broad range of psychological research domains, such as perception psychology, social and developmental psychology, cognitive psychology, and behavioral decision theory. The task of environmental psychology is still to combine these existing theories as well as to develop new ones for better understanding specific people–environment relations, and for supporting environmental designers and managers in enhancing the well-being of present and future generations.

Overall, present-day environmental psychology can be characterized not only as a highly socially relevant field of applied psychology, but also as an opportunity

for developing psychology in a more context-related direction, that is, toward that real world psychology continuously envisaged in environmental psychology from its beginnings to the present.

See Also the Following Articles

Environmental Stress ■ Intergroup Relations and Culture ■ Personal Space ■ Stress ■ Territoriality

Further Reading

Altman, I. (1988). Process, transactional/contextual and outcome research: An alternative to the traditional distinction between basic and applied research. *Social Behaviour*, *3*, 259–280.

Barker, G. A. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior.* Stanford, CA: Stanford University Press.

Bazerman, M. H., Messick, D. M., Tenbrunsel, A. E., & Wade-Benzoni, K. A. (Eds.) (1997). *Environment, ethics and behavior.* San Francisco: The New Lexington Press.

Bechtel, R. (1997). *Environment and behavior: An introduction.* London, UK: Sage.

Bechtel, R., & Churchman, A. (Eds.) (2002). *Handbook of environmental psychology.* New York: Wiley.

Bonnes, M., Lee, T., & Bonaiuto, M. (Eds.) (2003). *Psychological theories for environmental issues.* Aldershot, UK: Ashgate.

Bonnes, M., & Secchiaroli, G. (1995). *Environmental psychology. A psycho-social introduction.* London, UK: Sage.

Canter, D., & Lee, T. (Eds.) (1974). *Psychology and the built environment.* London, UK: Architectural Press.

Gifford, R. (2002). *Environmental psychology: Principles and practice* (3rd ed.). Collville, Canada: Optimal Books.

Kaplan, R., Kaplan, S., & Ryan, R. L. (1998). *With people in mind: Design and management of everyday nature.* Washington, D.C.: Island Press.

Levy-Leboyer, C. (1980). *Psychologie et environnement.* Paris: Presses Universitaires de France.

Stern, P., Young, O. R., & Druckman, D. (Eds.) (1992). *Global environmental change: Understanding the human dimensions.* Washington, D.C.: The National Academies Press.

Stokols, D., & Altman, I. (Eds.) (1987). *Handbook of environmental psychology.* New York: Wiley.

Stokols, D. (1978). Environmental psychology. *Annual Review of Psychology*, *29*, 253–295.

Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science*, *224*, 420–421.

Uzzell, D. (2003). Our uncommon future. In R. Garcia-Mira, J. M. Sabucedo Cameselle, & J. Romay Martinez (Eds.), *Culture, environmental action and sustainability* (pp. 21–39). Gottingen: Hogrefe & Huber.

Van Vugt, M., Snyder, M., Tyler, T. R., & Biel, A. (2000). *Cooperation in modern societies.* London, UK: Routledge.

Wapner, S., Demick, J., & Yamamoto, T., Minami, H. (2000). *Theoretical perspectives in environment-behavior research.* New York: Kluwer Academic.

Environmental Stress

Gary W. Evans
Cornell University, Ithaca, New York, USA

Sheldon Cohen
Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

GLOSSARY

adaptation level shifts in environmental perception History of experience with an environmental stressor changes referential criteria for environmental perception of current environmental conditions.

behavioral aftereffects Performance deficits following exposure to environmental stressors.

cognitive consequences of learned helplessness Lower mastery, reduced self-efficacy, or greater externality in locus of control from chronic environmental stressor exposure.

coping perseveration The use of well-learned coping strategies when an environmental condition is no longer present.

cumulative fatigue The buildup of fatigue from expenditure of energy to cope with an environmental stressor; the three types of cumulative fatigue are behavioral aftereffects, spillover, and diminished coping with multiple stressors.

diminished coping with multiple stressors The reduced ability to cope with a subsequent stressor following exposure to an environmental stressor.

effort by environmental stressor interaction Physiological mobilization produced by cognitive or physical effort to maintain task performance during stressor exposure.

environmental stressors Physical characteristics of the environment that produce stress (e.g., commuting, crowding, noise, pollution).

learned helplessness The realization, caused by uncontrollable environmental events, that one cannot control one's environment. This may result in diminished motivation to assert control even when feasible to do so, the inability to learn that subsequent challenges that are objectively controllable can be controlled, and feelings of anxiety and depression.

overgeneralization Strategies for coping with chronic environmental stressors may become overlearned so that coping perseveration occurs in the absence of the stressor; adaptation level shifts in environmental perception may also occur.

physiological mobilization Elevated physiological stress produced by chronic exposure to environmental stressors; also caused by acute environmental stressor exposure in combination with the effort to maintain optimum task performance.

spillover The negative affect, strained interpersonal relationships, or fatigue produced by exposure to an environmental stressor in one setting that carries over into another setting.

Apart from the direct physical insults produced by toxins, mental health, psychophysiology, performance, and human motivation are all affected by environmental quality. In this article, we present a theoretical framework for understanding how environmental quality impacts human behavior. Human beings are not passive recipients of environmental conditions; instead, they work to optimize the balance between environmental conditions and human needs. Human beings thrive in a vast array of ecological niches.

However, we pay a price for our enormous adaptational capacities. The adaptive costs of coping with suboptimal physical conditions provide a theoretical framework for describing how the physical environment influences human behavior. We develop this theoretical framework and illustrate psychological impacts of poor environmental quality.

1. THEORETICAL FRAMEWORK

There are four categories of the adaptive costs of coping with poor environmental quality that have psychological consequences. These categories are cumulative fatigue, learned helplessness, physiological mobilization, and overgeneralization.

1.1. Cumulative Fatigue

It takes energy to cope with suboptimal environmental conditions. Expenditure of energy, particularly when demands are high or prolonged, causes fatigue. Three types of cumulative fatigue are behavioral aftereffects, spillover, and diminished coping with multiple stressors.

1.1.1. Behavioral Aftereffects

Behavioral aftereffects refer to impaired task performance following exposure to an environmental demand. In a typical aftereffects paradigm, an individual works on a task during a physical stressor such as noise. Following cessation of stressor exposure, the individual is asked to perform another task. Compromised performance on the second task, when the stressor is no longer present, is an index of behavioral aftereffects.

In the Glass and Singer behavioral aftereffects paradigm, subjects work on a simple task while exposed to an environmental stressor. Environmental stressors that have been studied include noise, crowding, temperature, and air pollution. During the environmental exposure period, subjects worked on simple tasks (e.g., anagrams). Performance during exposure is generally unaffected. Immediately following task performance under the environmental stressor, the stressor is terminated, and the subject is given a behavioral aftereffects task. Two commonly used aftereffect measures are persistence on a difficult or unsolvable puzzle and ability to detect errors in a proofreading task. Table I depicts a typical set of findings from one of Glass and Singer's studies, in which immediately after uncontrollable noise exposure was terminated, subjects were

TABLE I
Glass and Singer Aftereffects Data Following Uncontrollable Noise

	Noise	Quiet
Number of attempts on two unsolvable puzzles	10.88	41.20

administered a task that assessed their persistence on unsolvable puzzles.

In addition to the large number of replications of the Glass and Singer aftereffects data across a wide array of environmental stressors, another aspect of this paradigm is noteworthy. Several investigators have examined behavioral aftereffects of chronic environmental stressors such as commuting, crowding, and noise. Bullinger *et al.* noted that the longer the duration of noise exposure in the case of the opening of a new airport, the stronger were the impacts of chronic noise exposure on behavioral aftereffects (see Fig. 1).

1.1.2. Spillover

A second type of cumulative fatigue produced by poor environmental conditions is spillover. Spillover occurs when conditions in one setting influence a subject's well being in another location. A common example of spillover is when poor working conditions interfere with
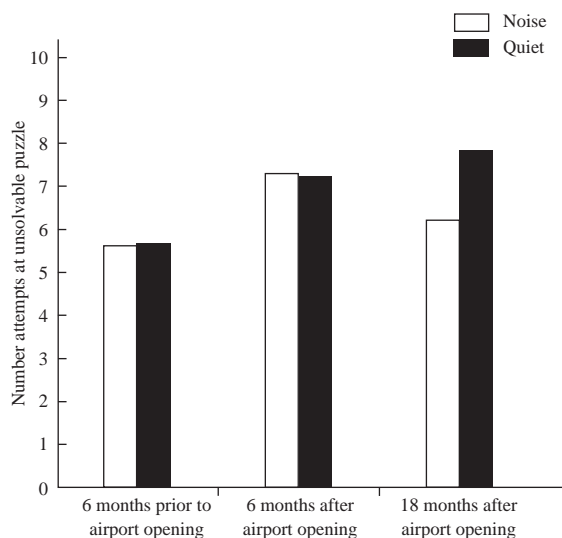

FIGURE 1 Task persistence among elementary aged school children in relation to airport noise exposure before and after the opening of a new airport. Controls matched on SES.

home life. A more congested commute creates interpersonal strains with family members at home. Epstein and Karlin reported that men but not women following a short term crowding experience in the laboratory felt less cohesive and more competitive. Immediately following exposure to crowded shopping conditions or high noise levels in a laboratory, individuals are less likely to behave altruistically when given a chance to provide help in a low stress setting. Table II provides illustrative data. A stranger (experimental confederate) appeared to be looking for a dropped contact lens on the floor of a shopping center in a low-density, inactive area of the center. Immediately prior to the assessment of spillover, subjects had been randomly assigned to perform a series of shopping tasks in noisy and crowded or relatively desolate areas of a mall.

1.1.3. Diminished Coping with Multiple Stressors

The effort applied in coping with environmental stressors can result in reduced capacity for coping with subsequent demands. Reduced coping capacity has been found following exposure to noise, crowding, high temperature, and pollutants. For example, the adverse effects of traffic noise on blood pressure in school children are accentuated if they also live in noisy homes. Residents of noisy areas respond more negatively to stressful events than residents with similar backgrounds (SES) who live in quiet areas. Crowding exacerbates the impact of high temperature on negative affect. Maxwell noted that the adverse influence of high-density daycare conditions on children's levels of behavioral disturbances was accentuated by crowding levels in the child's home. Researchers have also investigated crowding interactions with psychosocial rather than physical stressors. For example, elementary school aged children experienced greater physiological stress and psychological

distress in relation to greater family turmoil if they lived in more crowded homes and the effects of daily hassles on psychological health in adults were accentuated in high density homes.

Air pollution also accentuates adverse responses to other stressors. The effects of polluted working conditions on respiratory and dermatological symptoms were greater for workers also contending with higher levels of job stress. Evans and colleagues found that ozone was linked to psychological distress, but only among individuals also experiencing stressful life events (see Fig. 2).

1.2. Learned Helplessness

Attempts to cope with uncontrollable environments can result in learned helplessness. Individuals who try to abate a negative environmental condition and are unable to do so eventually learn that their efforts to control their environment are fruitless. They become unmotivated to assert control (motivational consequences), even when it is feasible to do so, unable to learn that subsequent challenges they confront that are objectively controllable can be controlled (cognitive consequences), and become anxious and depressed (affective consequences).

1.2.1. Cognitive Consequences

Experiencing uncontrollable environments can lead to the inability to learn that controllable environments

---

TABLE II
Altruistic Behavior Following a Simulated Shopping Task in a Crowded or Uncrowded Shopping Center

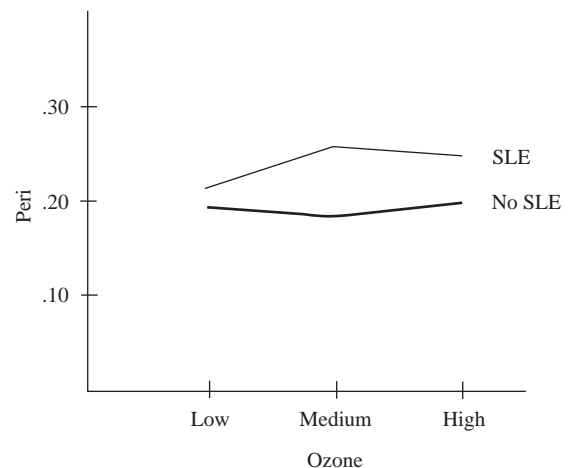|	*Crowded*	*Uncrowded*
Percent of subjects who helped	16.5	56.5
Seconds spent looking for contact lens	15	41

FIGURE 2 Psychological distress as a function of the interaction of exposure to pollution and recent stressful life events among a representative sample in Los Angeles.

can be managed or affected by personal effort. For example, individuals living under crowded conditions report less perceived control or feelings of mastery over their immediate surroundings compared to others living under less crowded conditions. Traffic congestion levels negatively impact feelings of control over commuting itself and in general among car drivers. Bus drivers facing more traffic congestion report significantly less control over their job.

Community airport noise exposure has also been linked to diminished feelings of personal control. For example, the typical maturational trajectory of increased internal locus of control was altered among elementary school children living for longer periods of time in airport noise impact zones (see Table III). SES controls were incorporated.

1.2.2. Motivational Consequences

As discussed previously, exposure to uncontrollable noise, crowding, or air pollution leads to deficits in behavioral aftereffects tasks. In the original aftereffects studies, Glass and Singer included an independent variable that points toward learned helplessness as a possible mechanism for behavioral aftereffects. When subjects were led to believe they could terminate the noise if it became too aversive by pushing a button, there were no behavioral aftereffects of noise. This occurred even though no one in this "perceived control" condition actually pushed the button. In short, aftereffects occurred only among those exposed to "uncontrollable" stressors. Table IV adds the noise with perceived control condition to our original behavioral aftereffects example shown before in Table I. Subjects were randomly assigned to conditions.

This salutary effect of perceived control over noise on behavioral aftereffects has been widely replicated. Parallel results have been uncovered in studies of crowding and pollution.

TABLE III
Chronic Noise Exposure and Control Beliefs among Elementary School Children[a]

	Years in residence	
	<3	>3
Low noise	5.56	6.00
High noise	6.00	5.43

[a] Higher numbers reflect greater internality.

TABLE IV
Aftereffects of Noise, Quiet, and Noise with Perceived Control

	Noise	Quiet	Noise with perceived control
Number of attempts on two unsolvable puzzles	10.88	41.20	41.67

More direct evidence that suboptimal environmental quality can produce helplessness comes from two paradigms. In one paradigm, participants performed a task in order to avoid a noxious stimulus. For some subjects, exposure to the noxious stimulus was contingent upon their performance—by performing well, they could minimize exposure to the noxious stimulus. Other subjects were "yoked" to the first group in that they were exposed to the same noxious stimulus for the same duration of time. For these subjects, however, exposure to the noxious stimulus was independent of their performance. These conditions (contingent vs noncontingent performance and punishment) were then followed by a second phase in which a performance contingent task was administered. For this task, avoidance of the noxious stimulus was possible for all subjects. Prior exposure to the inescapable stimulus condition induced learned helplessness in the second phase. Even when an instrumental response was available, subjects who had been in the initial inescapable condition took much longer to learn the escape response. Of particular interest is the fact that exposure to inescapable noise versus escapable noise reliably produces learned helplessness in adults.

In a second paradigm, vulnerability to the induction of helplessness is related to chronic environmental conditions. Cohen *et al.* gave elementary school children in either airport noise-impacted or quiet schools a jigsaw puzzle to work on. More children from high-noise areas failed to solve the puzzle. Of those children who failed, a larger percent of children from noise-exposed areas failed the puzzle because they simply gave up trying prior to the end of the testing period (see Table V). SES controls were included.

Crowded living conditions similarly affect learned helplessness in children. For example, the longer college students lived in a crowded dormitory, the greater their helplessness behaviors compared to students in

TABLE V

Percentage of Children Who Gave up on a Solvable Jigsaw Puzzle prior to the Total Time Available[a]

	Noisy school	*Quiet school*
Percent giving up	31	7

[a]Total time available was 4 minutes.

uncrowded dormitories. Moreover, these helplessness behaviors closely mirrored the extent to which they felt powerless to regulate social interaction in their dormitory.

1.3. Physiological Mobilization

The marshalling of coping resources requires energy. This expenditure of energy requires mobilization of physiological resources, as reflected by elevated physiological stress. This occurs primarily in cardiovascular and neuroendocrine systems. Mobilization of energy in response to acute environmental demands is adaptive. The problem occurs with chronic physiological mobilization in order to meet repeated environmental demands. These elevations produce wear and tear on the body, alter the body's ability to turn off these systems when external demands cease, and can eventually lead to physical and psychological morbidity.

Effort and physiological mobilization are interrelated. When performing a task under a stressor, if effort is high, task performance is maintained but physiological mobilization escalates. If effort is low under a stressor, performance suffers, but physiological mobilization is minimal. Thus, there is an effort by stressor tradeoff.

1.3.1. Physiological Stress

A large number of studies have been conducted on occupational noise exposure and blood pressure. The data are mixed and methodological flaws plague many of these studies. Community noise studies have used better designs and revealed consistent elevations in blood pressure in children exposed to airport noise. Fig. 3 shows data comparing school children's physiological stress levels before and after the opening of an airport. Children were tested 1 year before the airport opened, 1 year after, and then again 6 months later.

Crowding has similar effects on the physiological stress response. Crowding has been associated with elevated blood pressure in the laboratory and among institutional populations and in some but not all studies of crowded housing and neighborhoods. Traffic congestion has been associated with elevated blood pressure and stress hormones; more crowded or more difficult train commuting elevates stress hormones as well.
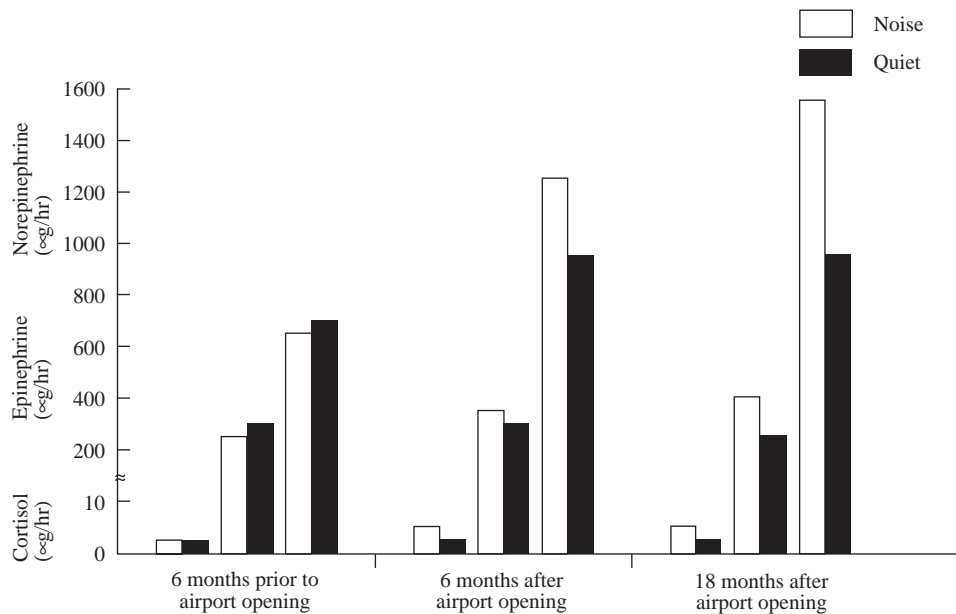

FIGURE 3 Physiological stress responses in children in a prospective study of airport noise. The samples were well matches on SES.

1.3.2. Effort by Stressor Tradeoff

Some of the physiological stress accompanying environmental stressors is caused by increased expenditure of effort to maintain task performance. Cardiovascular and neuroendocrine responses to acute noise rapidly habituate, returning to baseline levels. This habituation does not occur, however, when subjects simultaneously work on cognitively demanding tasks during noisy conditions. More direct evidence for the role of effort emanates from studies showing a performance–physiological stress tradeoff. For example, Tafalla and Evans manipulated cognitive effort and noise in a laboratory experiment, finding that noise adversely affected performance but had little impact on cortisol when effort was low. When participants worked under high effort on a task during noise, the opposite profile emerged: no performance effects of noise but elevated physiological stress. See Table VI for an illustration of some of their findings.

Field studies have uncovered analogous job demand by occupational noise interactions on physiological stress. In one study, the impact of noise on blood pressure more than doubled under high- vs low-demands jobs. A similar effort by stressor interaction on performance and physiology was found in a laboratory investigation of crowding.

1.4. Overgeneralization

Humans respond to environmental insults with strategies to optimize the balance between demands and resources. If the demands are chronic, however, coping strategies can become habitual, persevering even when they are not appropriate. Adaptation level shifts in environmental perception may also occur. This is when the reference criterion one uses to judge a current stimulus is influenced by prior history with that same stimulus. These shifts in referential criteria reflect habituation or reduce sensitivity to the stimulus.

1.4.1. Coping Perseveration

Children chronically exposed to noise appear to adopt a cognitive strategy of tuning or filtering out the noise. This strategy, while adaptive when noise is present, may be harmful if it becomes a characteristic strategy for processing auditory information, including meaningful stimuli such as speech. Cohen *et al.* showed that the louder the exposure to traffic noise at home, the more trouble children had in a task that assessed their abilities to discriminate between phonemes (e.g., the sound for b and d, or p and b). The children were prescreened for normal hearing thresholds. As shown in Table VII, noise exposure impairs phoneme perception, which in turn damages reading acquisition. SES controls were incorporated. Other noise studies have found similar relationships.

Perseveration of coping strategies has also been uncovered in studies of crowding. A common strategy for coping with crowding is social withdrawal. If people cope with overcrowded living conditions by socially withdrawing, an unintended consequence might be the breakdown of socially supportive relationships. Evans *et al.*, Lakey, and Lepore *et al.* found evidence for lower levels of perceived support in more crowded residences. The higher the ratio of people per room, independent of SES, the lower the levels of social support. These studies also showed that the link between residential crowding and poor psychological health was mediated by social support.

The adverse effects of residential crowding on social support appear to be caused by social withdrawal. In one study, college students who lived in high- or low-density apartments were brought into the laboratory. The lab was not crowded. The subject and a

---

TABLE VI
Stressor by Effort Interaction on Physiological Stress and Task Performance

	Low effort	High effort
Quiet		
Reaction time	5138	4250
Cortisol	.022	.013
Noise		
Reaction time	8110	5987
Cortisol	.024	.026

TABLE VII
Chronic Noise Exposure, Auditory Discrimination, and Reading Acquisition

Variable	Variance accounted for (%)
Noise	Phoneme perception = 19
Noise	Reading = 4
Noise, after statistically partialling out phoneme perception	Reading = 0

confederate were placed into situations in which they each needed social support at different times. As shown in Fig. 4, subjects who lived in crowded homes were significantly more unresponsive to the confederate's offers of support. The obverse occurred as well.

1.4.2. Adaptation Level Shifts

People may become so accustomed to suboptimal environmental conditions that they become desensitized to poor environmental quality. For example, an auditory stimulus will be perceived as softer if loud sounds have recently been experienced. Berglund *et al.* found that a given level of noise is less annoying if one has a history of living in a noisy area. Parallel results have been found for chronic residential crowding history and judgments of perceived crowding.

Wohlwill and Kohn examined residential histories and judgments of environmental quality. Adults who had recently migrated to a medium-sized community (Harrisburg, Pennsylvania) were asked to judge the attributes of their new community. Migrants from urban areas judged Harrisburg as significantly less noisy, less crowded, and less polluted than did their counterparts from small towns. In a second study, photographs of environmental qualities of communities varying in scale were shown. Adults from small towns consistently perceived greater levels of noise and crowding for a given photograph compared to individuals from large metropolitan areas.

People's judgments of perceived air quality habituate with continued experience of pollution. Flachsbart and Phillips showed that smaller changes in smog are
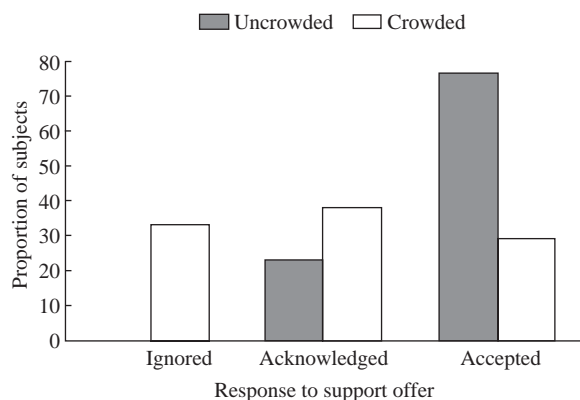
FIGURE 4 Proportion of residents from crowded and uncrowded apartments who ignored, acknowledged, or accepted confederate's social support.

TABLE VIII

Signal Detection Data for Recognition of Smog for People with Long-Term or Little Prior Experience with Air Pollution

History of air pollution	Response criterion	Detection sensitivity
Long-term experience		
Scene 1	.709	2.171
Scene 2	.637	2.052
Little prior experience		
Scene 1	.211	2.751
Scene 2	.276	2.340

required to alter judgments of perceived air quality for those unaccustomed to poor air quality compared to those who have prior familiarity with smog. Evans *et al.* analyzed pollution perception with signal detection theory. Signal detection theory enables one to separate perceptual processes from response criteria when judging stimuli. For example, if one is expecting an important phone call, one's threshold (response criteria) for responding to auditory stimuli shifts. It takes less of an indication that a bell has rung under these circumstances in order for someone to check the phone. Note that actual detection sensitivity to hear a bell is not changed, just response criterion.

College students who had recently migrated to southern California were tested. The students were well matched except with respect to prior histories of experience with air pollution. The students were shown two different, unfamiliar scenes of a distant vista. Each scene was shown multiple times with varying degrees of photochemical smog present. Subjects reported whether they saw smog and rated their confidence in their judgments. Students with prior histories of air pollution were less likely to report that pollution was present in a scene compared to people relatively unfamiliar with pollution (Table VIII). Actual visual thresholds of seeing air pollution (detection sensitivity) were equivalent.

2. CONCLUSION

The adaptive capabilities of human beings enables us to cope with a wide range of environmental conditions. Table IX provides a summary of some of the psychological consequences of human attempts to cope with suboptimal environmental quality.

TABLE IX
Adaptive Costs of Coping with Suboptimal Environmental Conditions

Category of adaptive cost	Types of effects
Cumulative fatigue	*Behavioral aftereffects:* Performance deficits following exposure to environmental stressors
	Spillover: Negative affect, strained interpersonal relationships, or fatigue produced by exposure to an environmental stressor in a different setting
	Diminished coping with multiple stressors: Reduced ability to cope with a subsequent stressor following exposure to an environmental stressor
Learned helplessness	*Cognitive consequences:* Lower mastery, reduced self-efficacy, greater externality in locus of control from chronic environmental stressor exposure
	Motivational consequences: Diminished motivation in performance or inability to learn new tasks produced by exposure to uncontrollable environmental stressors
Physiological mobilization	*Physiological stress:* Elevated cardiovascular and neuroendocrine activity following stressor exposure
	Effort by stressor tradeoff: Physiological mobilization by cognitive or physical effort to maintain task performance during stressor exposure
Overgeneralization	*Coping perseveration:* Use of well-learned coping strategies when environmental condition no longer present
	Adaptation level shifts: History of experience with an environmental stressor changes referential criteria for environmental perception of current environmental conditions

Acknowledgments

See Also the Following Articles

Coping ▪ Privacy ▪ Stress

Further Reading

Baum, A., Gatchel, R. J., Aiello, J. R., & Thompson, D. E. (1981). Cognitive mediation of environmental stress. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 513–533). Hillsdale, NJ: Erlbaum.

Berglund, B., Berglund, U., & Lindvall, T. (1975). A study of response criteria in populations exposed to aircraft noise. *Journal of Sound and Vibration, 41*, 33–39.

Berglund, B., & Lindvall, T. (1995). Community noise. *Archives of the Center for Sensory Research, 2*, 1–195.

Booth, A. (1976). *Urban crowding and its consequences*. New York: Praeger.

Bullinger, M., Hygge, S., Evans, G. W., Meis, M., & van Mackensen, S. (1999). The psychological cost of aircraft noise for children. *Zentralblatt fur Hygeine und Umweltmedizin, 202*, 127–138.

Cohen, S. (1980). Aftereffects of stress on human performance and social behavior: A review of research and theory. *Psychological Bulletin, 88*, 82–108.

Cohen, S., Evans, G. W., Stokols, D., & Krantz, D. S. (1986). *Behavior, health, and environmental stress*. NY: Plenum.

Cohen, S., Glass, D. C., & Singer, J. E. (1973). Apartment noise, auditory discrimination, and reading ability in children. *Journal of Experimental Social Psychology, 9*, 407–422.

Cohen, S., & Spacapan, S. (1978). The aftereffects of stress: An attentional interpretation. *Environmental Psychology and Nonverbal Behavior, 3*, 43–57.

Cottington, E. M., Matthews, K. A., Talbott, E., & Kuller, L. (1983). *Occupational stress and diastolic blood pressure in a blue collar population: The Pittsburgh Noise Hypertension Project*. Annual Meeting of the Society for Epidemiological Research. Winnipeg.

D'Atri, D. A. (1975). Psychophysiological responses to crowding. *Environment and Behavior, 7*, 237–250.

Dubos, R. (1965). *Man adapting*. New Haven, CT: Yale University Press.

Eoyang, C. K. (1974). Effects of group size and privacy in residential crowding. *Journal of Personality and Social Psychology, 30*, 389–392.

Epstein, Y. M., & Karlin, R. A. (1975). Effects of acute experimental crowding. *Journal of Applied Social Psychology*, 5, 34–53.

Epstein, Y. M., Lehrer, P., & Woolfolk, R. (1978). *Physiological, cognitive, and behavioral effects of repeated exposure to crowding*. Washington, DC: American Psychological Association.

Evans, G. W. (1979). Behavioral and physiological consequences of crowding in humans. *Journal of Applied Social Psychology*, 9, 27–46.

Evans, G. W. (2001). Environmental stress and health. In A. Baum, T. A. Revenson, & J. E. Singer (Eds.), *Handbook of health psychology* (pp. 365–385). Mahwah, NJ: Erlbaum.

Evans, G. W., Bullinger, M., & Hygge, S. (1998). Chronic noise exposure and physiological response: A prospective, longitudinal study of children under environmental stress. *Psychological Science*, 9, 75–77.

Evans, G. W., & Carrere, S. (1991). Traffic congestion, perceived control, and psychophysiological stress among urban bus drivers. *Journal of Applied Psychology*, 76, 658–663.

Evans, G. W., Hygge, S., & Bullinger, M. (1995). Chronic noise and psychological stress. *Psychological Science*, 6, 333–338.

Evans, G. W., Jacobs, S. V., & Frager, N. B. (1982). Adaptation to air pollution. *Journal of Environmental Psychology*, 2, 99–108.

Evans, G. W., Jacobs, S. V., Dooley, D., & Catalano, R. (1987). The interaction of stressful life events and chronic strains on community mental health. *American Journal of Community Psychology*, 15, 23–34.

Evans, G. W., & Lepore, S. J. (1993). Household crowding and social support: A quasi-experimental analysis. *Journal of Personality and Social Psychology*, 65, 308–316.

Evans, G. W., Lepore, S. J., Shejwal, B. F., & Palsane, M. N. (1998). Chronic residential crowding and children's well being: An ecological perspective. *Child Development*, 69, 1514–1523.

Evans, G. W., & Maxwell, L. E. (1997). Chronic noise exposure and reading deficits: The mediating effects of language acquisition. *Environment and Behavior*, 29, 710–728.

Evans, G. W., Palsane, M. N., Lepore, S. J., & Martin, J. (1989). Residential density and psychological health: The mediating effects of social support. *Journal of Personality and Social Psychology*, 57, 994–999.

Evans, G. W., & Saegert, S. (2000). Residential crowding in the context of inner city poverty. In S. Wapner, J. Demick, H. Minami, & T. Yamamoto (Eds.), *Theoretical perspectives in environment behavior research: Underlying assumptions, research problems, and relationships* (pp. 247–268). New York: Plenum.

Evans, G. W., Saegert, S., & Harris, R. (2001). Residential density and psychological health among children in low income families. *Environment and Behavior*, 33, 165–180.

Evans, G. W., & Stecker, R. (2004). Motivational consequences of environmental stress. *Journal of Environmental Psychology,* in press.

Flaschsbart, P. G., & Phillips, S. (1980). An index and model of human response to air quality. *Journal of the Air Pollution Control Association*, 30, 759–768.

Fleming, I., Baum, A., & Weiss, L. (1987). Social density and perceived control as mediators of crowding stress in high density residential neighborhoods. *Journal of Personality and Social Psychology*, 52, 899–906.

Glass, D. C., & Singer, J. E. (1972). *Urban stress*. New York: Academic Press.

Griffitt, W., & Veitch, R. (1971). Influences of population density on interpersonal affective behavior. *Journal of Personality and Social Psychology*, 17, 92–98.

Guski, R., & Rohrman, B. (1981). Psychological aspects of environmental noise. *Journal of Environmental Policy*, 2, 183–212.

Helson, H. (1964). *Adaptation level theory*. New York: Harper & Row.

Herronkohl, R., & Egoff, B. (1974). *Perceived crowding, familiarity, and the choice of a residence*. American Psychological Association. New Orleans.

House, J. S., Mc Michael, A., Wells, J., Kaplan, B., & Landerman, L. (1979). Occupational stress and health among factory workers. *Journal of Health and Social Behavior*, 20, 139–160.

Kluger, A. (1998). Commute variability and strain. *Journal of Occupational Behavior*, 19, 147–165.

Kryter, K. (1994). *The handbook of hearing and the effects of noise*. New York: Academic Press.

Lakey, B. (1989). Personal and environmental antecedents of perceived social support developed at college. *American Journal of Community Psychology*, 17, 503–519.

Lepore, S. J., Evans, G. W., & Palsane, M. N. (1991). Social hassles and psychological health in the context of crowding. *Journal of Health and Social Behavior*, 32, 357–367.

Lepore, S. J., Evans, G. W., & Schneider, M. (1991). The dynamic role of social support in the link between chronic stress and psychological distress. *Journal of Personality and Social Psychology*, 61, 899–909.

Lundberg, U. (1976). Urban commuting: Crowdedness and catecholamine excretion. *Journal of Human Stress*, 2, 26–34.

National Research Council. (1991). *Environmental epidemiology*, Vol. 1. Washington, DC: National Academy of Sciences.

Maxwell, L. E. (1996). Multiple effects of home and day care crowding. *Environment and Behavior*, 28, 494–511.

Maxwell, L. E., & Evans, G. W. (2000). The effects of noise on preschool children's prereading skills. *Journal of Environmental Psychology*, 20, 91–97.

Mc Callum, R., Rusbult, C., Hong, G., Walden, T., & Schopler, J. (1979). Effects of resource availability and importance of behavior on the experience of crowding. *Journal of Personality and Social Psychology*, 37, 1304–1313.

Mc Ewen, B. S. (2000). The neurobiology of stress: From serendipity to clinical relevance. *Brain Research*, 886, 172–189.

Medical Research Council. (1997). *The nonauditory effects of noise*. Report R10, Institute for Environment and Health. Leicester, UK: University of Leicester.

Melamed, S., Fried, Y., & Froom, P. (2001). The interactive effect of chronic exposure to noise and job complexity on changes in blood pressure and job satisfaction: A longitudinal study of industrial employees. *Journal of Occupational Health Psychology*, 6, 182–195.

Moch-Sibony, A. (1984). Study of the effects of noise on the personality and certain psychomotor and intellectual aspects of children after a prolonged exposure. *Travail Humane*, 47, 155–165.

Novaco, R. W., Kliewer, W., & Broquet, A. (1991). Home environmental consequences of commute travel impedance. *American Journal of Community Psychology*, 19, 881–909.

Novaco, R. W., Stokols, D., Campbell, J., & Stokols, J. (1979). Transportation stress, and community psychology. *American Journal of Community Psychology*, 7, 361–380.

Novaco, R. W., Stokols, D., & Milanesi, L. (1990). Objective and subjective dimensions of travel impedance as determinants of commuting stress. *American Journal of Community Psychology*, 18, 231–257.

Paulus, P. B. (1988). *Prison crowding: A psychological perspective*. New York: Springer-Verlag.

Peterson, C., Maier, S., & Seligman, M. E. P. (1993). *Learned helplessness*. New York: Oxford.

Regoceva, V., & Kellcrova, E. (1995). Effects of urban noise pollution on blood pressure and heart rate in school children. *Journal of Hypertension*, 13, 408–412.

Repetti, R. L. (1993). The effect of workload and the social environment at work on health. In L. Goldberger, & Breznitz, S. (Eds.), *Handbook of stress,* 2nd ed. (pp. 368–385). New York: Free Press.

Rodin, J. (1976). Density, perceived choice, and response to controllable and uncontrollable outcomes. *Journal of Experimental Social Psychology*, 12, 564–578.

Rotton, J. (1983). Affective and cognitive consequences of malodorous pollution. *Basic and Applied Social Psychology*, 4, 171–191.

Ruback, R. B., & Pandey, J. (1991). Crowding, perceived control and relative power: An analysis of households in India. *Journal of Applied Social Psychology*, 21, 315–344.

Ruback, R. B., & Pandey, J. (1992). Very hot and really crowded. *Environment and Behavior*, 24, 527–554.

Rydstedt, L. W., Johansson, G., & Evans, G. W. (1998). The human side of the road: Improving the working conditions of urban bus drivers. *Journal of Occupational Health Psychology*, 3, 161–171.

Saegert, S. (1978). High density environments: Their personal and social consequences. In A. Baum, & Y. M. Epstein (Eds.), *Human response to crowding* (pp. 259–282). Hillsdale, NJ: Erlbaum.

Schaeffer, M. A., Street, S., Singer, J. E., & Baum, A. (1988a). Effects of control on the stress reactions of commuters. *Journal of Applied Social Psychology*, 11, 944–957.

Sherrod, D. R. (1974). Crowding, perceived control, and behavioral aftereffects. *Journal of Applied Social Psychology*, 4, 171–186.

Sherrod, D. R., & Downs, R. (1974). Environmental determinants of altruism: The effects of stimulus overload and perceived control on helping. *Journal of Experimental Social Psychology*, 10, 468–479.

Stokols, D., Novaco, R. W., Campbell, J., & Stokols, J. (1978). Traffic congestion, Type A behavior, and stress. *Journal of Applied Psychology*, 63, 467–480.

Tafalla, R. J., & Evans, G. W. (1997). Noise, physiology, and human performance: The potential role of effort. *Journal of Occupational Health Psychology*, 2, 148–155.

van Kamp, I. (1990). *Coping with noise and its health consequences*. Groningen, Netherlands: Styx.

Welch, B. L. (1979). *Extra auditory health effects of industrial noise: Survey of foreign literature*. AHRL-TR-79–41. Aerospace Medical Research Laboratory. Dayton, OH: Wright Patterson Air Force Base.

Wener, R., Evans, G. W., Phillips, D., & Nadler, N. (2003). Running for the 7:45: The effects of public transit improvements in commuter stress. *Transportation*, 30, 203–220.

White, S., & Rotton, J. (1998). Type of commute, behavioral aftereffects, and cardio-vascular activity. *Environment and Behavior*, 30, 763–780.

Wohlwill, J. F., & Kohn, I. (1973). The environment as experienced by the migrant: An adaptation level view. *Representative Research in Social Psychology*, 4, 135–164.

Wohlwill, J. F., & Kohn, I. (1976). Dimensionalizing the environmental manifold. In S. Wapner, S. Cohen, & B. Kaplan (Eds.), *Experiencing the environment* (pp. 19–53). New York: Plenum.

Environmental Versus Individual Risk Taking: Perception, Decision, Behavior

Charles Vlek

University of Groningen, Groningen, The Netherlands

GLOSSARY

commons dilemma The (un)balance between individual interests and collective goods, depending on the discounting of temporally, spatially, and/or socially distant decision consequences of those involved.

coping appraisal Overall judgment of one's potential to handle an external threat, in terms of response efficacy, self-efficacy, and costs and barriers of action.

external versus internal control Events may occur through (external) "acts of god," (internal) human factors, or both. This underlies the difference between gambling tasks and control tasks.

frequency versus scenario information Historical relative frequencies versus future possible ways of accident occurrence, as bases for probability estimation.

judgmental heuristics and biases Simplified ways and tendencies to estimate probabilities and other uncertain quantities.

perceived risk Personal judgment, on various grounds, of the nature and seriousness of risk, often contrasted with expert-assessed or calculated risk based on scientific models and/or databases.

precautionary principle The willingness to decide for (costly) protective measures or against risky activity if there is reasonable suspicion of serious threat.

risk The possibility of a negative event (harm, loss, accident, disease, or disaster).

risk attitude The tendency either to venture something or not in view of the safe(r) status quo, depending on utilities of possible decision consequences.

risk homeostasis A risk taker's tendency to use externally increased safety for enhanced benefit achievement, thereby maintaining or returning to the original (target) risk level.

threat appraisal Overall judgment of an external threat, in terms of perceived seriousness, vulnerability, and associated benefits.

utility function The formal relationship between personal utility and the objective value of decision consequence variables (e.g., monetary outcome, driving speed, number of fatalities).

Managing risks is an everyday task of life, appearing in different scopes and scales. To understand environmental

risk management, we must be familiar with individual risk taking. In this article, basic concepts and theories about risk are reviewed, along with empirical findings about judgmental and behavioral mechanisms. For risk taking, perceived control often is more important than calculated gambling. The major dimensions of perceived technological risk are the undesirability of the consequences, perceived controllability, and importance of associated benefits. Collective risks often arise from negative external effects of numerous individual activities, such as driving a car. Several factors of collective rationality (or cooperation) are discussed, along with various intervention strategies for achieving a sustainable balance in such commons dilemmas. Conclusions focus on psychologically enlightened safety management, both small-scale and large-scale.

1. RISK LEVELS, DEFINITIONS, AND COMPLICATIONS

On December 9, 2003, four Latvian mountaineers were climbing Mount Cook, the highest peak of New Zealand. To ensure each other's safety, they were all connected to the same rope. Suddenly, one of them slipped on the icy slope. The other three at that moment were not standing firm enough to absorb the shock of their falling colleague. Unfortunately, the four sportsmen fell 300 meters and died on the spot.

In November 2003, more than 200 people died, an entire village was destroyed, and much farmland was turned useless by abundant river floods in the Northern Sumatran district of Lankat (Indonesia) following days of heavy monsoon rainfall. Wholesale deforestation, undertaken for the benefit of logging companies, appeared to be the major cause; it had been unofficially tolerated for many years despite the warnings from environmental protection groups and local communities.

1.1. Research Origins

Things can always go wrong. However promising or well-intended, any activity or situation may result in loss, harm, accident, or disaster (henceforth referred to as loss). Sometimes we are personally responsible and have to bear the damage ourselves. In other cases, risks are collective, arising through natural events, through the actions of others, or through both.

Statistics and economics have long dominated the thinking of researchers and policymakers about risk.

Mathematical statistics has laid the grounds for probabilistic risk analysis in the engineering sciences. Some 25 years ago, this elicited psychometric research on perceived risks of technologies. Neoclassical economic theory has evoked 40-year-old behavioral decision research, in which psychologists gradually developed authentic concepts and theories of their own. Social psychologists increasingly study commons dilemmas, capturing the tension between individual interests and collective goods, such as environmental sustainability.

Here, an overview of concepts, models, and findings on human risk judgment and decision making about safety is provided. The emphasis is on conceptual clarification and interrelation from different research directions. First considered is the concept of risk and individual risk taking, then the evaluation and management of collective environmental risks.

1.2. Different Levels of Risk

In risk taking there seem to be two extremes. On the small-scale side, the quality of a challenging activity, such as mountain climbing, car driving, or tight-rope walking across Niagara Falls, may actually suffer from the risk awareness of the participant, since that awareness might distract attention and raise worries, whereas the self-confident subject tries to perform optimally and prevent accidents from occurring. On the large-scale side, people may feel uneasy or threatened by environmental conditions such as ambient air pollution, street crime, or the chance of flooding. Such situations stimulate explicit risk assessment and the design of coordinated safety strategies, often involving many people.

For the moment, loosely defining risk as the possibility of undesirable consequences of an activity or situation (henceforth: activity), risks may be categorized as occurring on seven different levels of complexity, as shown in Table I, together with illustrative examples.

Table I shows varieties of risk associated with single, low-level activities that are normally accompanied by a relative unawareness of risk; safety is warranted by due attention, motivation, and experienced behavioral control. Table I also gives examples of worrisome risks whose cause or (more often) effective control is not sufficiently clear. One may roughly distinguish three types of risks: clearly perceptible risks (as in car traffic), scientifically clear but not publicly well-known risks (as from soil pollution), and new, unknown, or ambiguous risks (as from genetic engineering) that may call for precautionary measures.

TABLE I
Different Scale Levels of Risk and Risk Management

Level	Key words to illustrate risk (as "possible loss from …")
Personal	Smoking, drinking, illness, sports
Indoor	Air quality, radon, do-it-yourself jobs
Local	Traffic, toxic waste, urban smog
Regional	Air, soil, and groundwater pollution; GMO crops
Fluvial	Pollution of rivers, lakes, coastal zones; flooding
Continental	Acid precipitation, deforestation, glacier melting
Global	Ozone-layer depletion, climate change, sea level rise

1.3. Definitions of Risk

Following traditional statistics, risk is the probability of loss, where probability is estimated on the basis of long-term observed relative frequency of comparable losses. In business economics, in decision theory, and in practical safety management, risk also has other, both extended and more subjective definitions. Table II lists six formal definitions of risk that may be used for quantitative estimation.

The psychological problem with any definition in Table II is that both probability and size of loss are essentially judgmental components. Probability may be objectified, but only under rather limiting conditions, such as, for example, the assumption that the relevant system will behave in the future as it did in the (observed) past and, obviously, that a sufficient

TABLE II
Formal Definitions of Risk

- Probability of specified loss
- Size of possible loss
- Product of probability and size of loss
- Variance of total distribution of losses and benefits of decision alternative
- Relative weight of possible loss versus comparable gain (following utility curve)
- Two-dimensional set of points indicating probabilities of possible different losses[a]

[a]This may characterize, e.g., a factory, household, or transport on a multi-annual basis.

number of observations is available to compute relative frequencies. Size of loss similarly presupposes a fair specification of what such loss really means and how it eventually all works out (or not). Multi-attributable notions of loss (as a function of, e.g., people killed and wounded, property damage, environmental pollution) do not easily fit into classical risk definitions.

On an aggregate level (Table I), the international debate on climate change illustrates the scientific difficulties of environmental risk assessment and management.

1.4. Development of Risk Psychology

In critical response to technological risk analysis, social scientists, particularly psychologists, have explored people's risk perceptions and behaviors. Thus, a body of theoretical and empirical knowledge has developed that represents nontechnical and noncomputational ways of conceptualizing and grasping risk as an essential component of human performance, deliberation, and emotionality.

One result of this development is that nowadays the objectification of technological risk has limited (only "bandwidth") potential and that the actual assessment, decision making, and control of external risks is to a large extent a procedural and often multi-stakeholder matter involving attention, effort, competence, and communication with respect to technical as well as behavioral and social components of risky activities. Many environmental risks came to be seen as collective effects generated by numerous individual activities in commons dilemma situations (see Section 5). One challenge for psychology is to expand from existing theory and methodology about small-scale risk taking toward a better understanding of human judgment and behavior *vis-à-vis* large-scale environmental risks.

2. MODELS AND MECHANISMS OF INDIVIDUAL RISK TAKING

This section reviews a number of theoretical models and issues, giving various insights into the way we judge risks and deal with risky situations.

2.1. Classical Decision Theory

Following classical economic theory, a simple choice between two gambles may be portrayed as in Fig. 1,
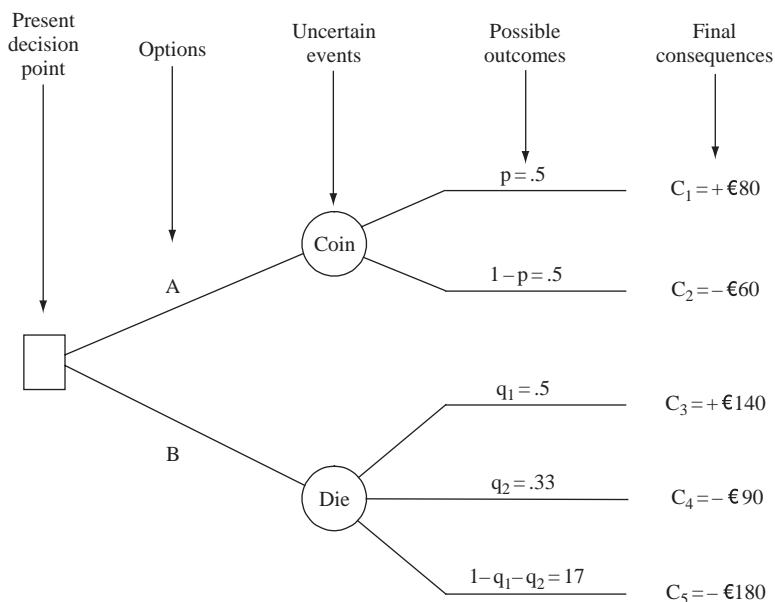
FIGURE 1 Decision tree representation of the choice between gambles A (flipping a coin) and B (throwing a die with possible outcomes 1–3, 4–5, or 6) and monetary consequences in euros.

which shows the perspective of a decision maker in terms of present options, possible outcomes of uncertain factors, and monetary consequences. Note that the expected value (EV) of both gambles A and B, that is, their long-term average return on repeated playing, amounts to €10. For a single play, which gamble would you prefer, and why? A rational EV maximizer's answer would be, "Either, because both gambles are equally worthy." EV theory, however, has been succeeded by expected utility (EU) theory, in which the subject's utility function for money is crucial. One traditional shape of such a curve is drawn in the right half of Fig. 2, which reflects the subject's supposed decreasing marginal utility of money. This would, for example, imply the subject's refusal to play a 50–50 bet for a fixed (sizeable) amount of money, since the possible loss weighs heavier than the comparable gain. Thus, in the example of Fig. 1, both gambles would be unattractive, but actually gamble A has a higher EU than gamble B.

As regards risk, in view of Table II, we may note that A and B have an equal (total) probability of loss, B has the highest possible size of loss, B has the highest (summed) product of probability and size of loss, and B has the highest outcome variance. So, gamble B is the most risky. One should realize, however, that not risk but EU is the rational basis of choice.
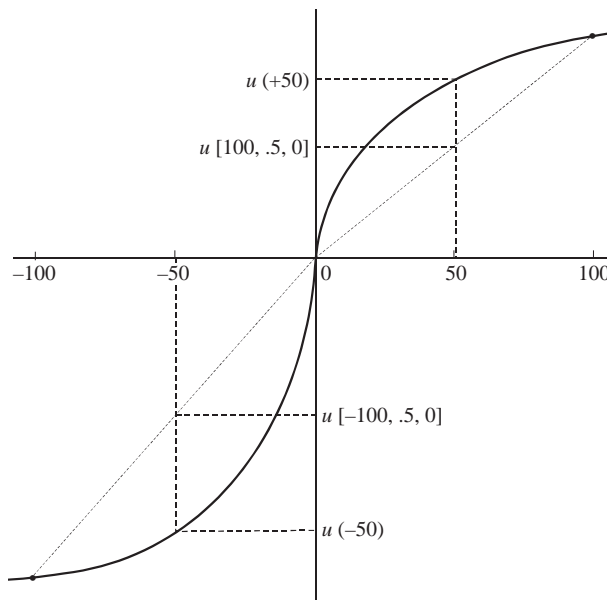

FIGURE 2 Two-component value function following prospect theory, indicating risk avoidance in the domain of gains and risk seeking in the domain of losses. Monetary amounts are indicated on the horizontal axis, utilities on the vertical one.

The negative acceleration of the utility curve in Fig. 2 fits the well-known Weber–Fechner law of psychophysics: as stimulus intensities increase, people become less sensitive to given intensity differences. For monetary amounts, the difference between 1020 and 1000 euro is psychologically less than the difference between 120 and 100 euro. This notion of decreasing marginal utility also applies in the negative (see Fig. 2). For example, the psychological difference between 120 and 100 (possible) accident fatalities is less than the difference between 20 and 0 fatalities.

2.2. Prospect Theory

The preceding reasoning is captured in a generalized value function following prospect theory as proposed by Kahneman and Tversky in 1979; see Fig. 2. Here, the utility function for losses is steeper than the utility function for gains: "losses loom larger than gains." This double utility function has special implications for gambling in the domains of (only) losses and gains. Suppose the subject is offered the choice between receiving €50 for sure and a 50–50 bet to win either €100 or nothing (written as [100, .50, 0]). Following the utility (u) curve for gains, $u(+50)$ is significantly higher than $u[100, .50, 0]$; see the "positive utility" axis in Fig. 2. Thus, the subject will avoid the risk of getting nothing and will prefer the sure amount of €50.

Now suppose that the same subject is offered the choice between a sure loss of €50 and a 50–50 bet to lose either €100 or €0 (written as [–100, .50, 0]). As the "negative utility" axis reveals, $u(–50)$ is significantly lower (more negative) than $u[–100, .50, 0]$. Hence, the subject will prefer the bet over the sure loss, thus choosing the risk of losing as much as €100, in the hopes of actually obtaining the "loss" of €0.

The main implications for risk taking of the double utility function represented in Fig. 2 are the following:

• Following Weber–Fechner, the greater the absolute value of two possible outcomes (losses or gains), the smaller a given difference between them is perceived.
• Losses weigh more heavily than do comparable gains (this means, for example, that a 50–50 bet of losing or winning €100 will be rejected).
• Subjects are risk seeking in the domain of losses but are risk averse in the domain of gains.
• The perception of gains and losses depends on the subject's reference point, which may be, for example, current assets, last year's earnings, maximum outcome, or "worst case."

From a rationalistic perspective, the last point is rather worrying. As Kahneman and Tversky and others have shown, smart reframing of a decision problem may induce a preference change in the subject, who thereby would behave inconsistently across different frames. Framing may also be conducted in other ways, as for example when any particular accident scenario or risk dimension is expressly presented in a conspicuous way in order to cause it to be more heavily weighted than other scenarios or dimensions.

2.3. Probability Estimation

Considering risk, probability is the likelihood of a future loss. Probability, however, is subject to different conditions, interpretations and estimation heuristics, so that numerical values may be of limited validity and reliability.

2.3.1. Probability Interpretations

Debates about risk may be complicated by three different interpretations of probability.

1. Frequentistic probability is the limit of a (past) relative frequency in a theoretically endless number of observations of an uncertain event (e.g., the flip of a coin or the crossing of a road junction) for a specified outcome (e.g., "tails" or a car accident) under substantially similar conditions.
2. Logical probability is a quantitative characteristic of a chance apparatus such as a coin, die, or wheel of fortune, which may be derived from its properties such as the number of possible outcomes, symmetry, and lack of bias.
3. Personalistic probability is someone's degree of belief in the future occurrence of an uncertain outcome (e.g., a nearby forest fire next year), which may be expressed in the subject's relative willingness, expressed in the odds proposed, to bet on it.

2.3.2. Estimation Heuristics

For practical risk assessment involving small likelihoods, probability estimation often is as crucial as it is difficult. Assessment difficulty may be due to

• incomplete or inconsistent information from various sources,
• unrepresentativeness of historical frequency data,

- impossibility or implausibility of logical system modeling (e.g., for industrial process failure), and
- lack of substantive knowledge, experience, and imagination as a basis for expert judgment (e.g., about system failure scenarios).

Moreover, in many practical situations, the probability of things going wrong is rather elusive, as, for example, in mountain climbing, road transport, and marine navigation. Thus, risk takers may use various mental tricks to estimate probabilities (when required) as a basis for risk comparisons, acceptance decisions, or safety maneuvers. Well-known probability heuristics are estimation by representativeness, cognitive availability, and anchoring and adjustment.

An event, for example, a particular series of sunny, cloudy, and rainy days, may be judged to be more probable when it is believed to be more representative of the underlying event-generating process, namely, the local climate.

Cognitive availability plays a role to the extent that examples may come to mind (e.g., financial losses, sad divorces, environmental accidents) that resemble the outcome whose probability is to be estimated. This availability heuristic also explains the great sensitivity and fear of car (or bus or train) passengers, especially non-drivers, who have actually experienced an unforgettable accident. "Never again" is often a motive for quickly enhanced safety measures in otherwise normal risk situations.

Anchoring and adjustment is an heuristic maneuver for coming up with a numerical probability estimate under a fair amount of second-order uncertainty. The idea here is that either a middle or an end scale value offers an anchor for an initial estimate, which is then conservatively adapted.

2.3.3. Frequency versus Scenario Information

Probability estimates are based on differing kinds of information: relative frequencies, logical system properties,

process similarities, remembered cases, and anchor values for adapted judgment. More-encompassing foundations for probability estimation are causal schemas, cognitive scripts, or mental images of future scenarios. The fundamental difference between frequency thinking and scenario thinking is indicated in the columns of Table III. Obviously, the frequentistic probability interpretation applies to the left column, while the personalistic (belief) interpretation applies to the right column. The logical probability interpretation may be applicable to both columns, in order to compute probability estimates without performing repeated observations (left column) or after some logical modeling of a multi-component system or process (right column), such as, for example, a petrochemical plant. Experimental subjects appeared to be sensitive to both frequency and scenario information. However, in a combined presentation of frequency and scenario information, the role of frequency information appears to be significantly suppressed. This suggests that for risk estimation, future orientation is more important than looking at performance history.

Remarkably enough, many great accidents or disasters have actually occurred because they had not occurred before (frequency information was lacking) and because responsible authorities did not believe them to be possible (there were no plausible scenarios) and thus failed to take the necessary precautions.

2.3.4. External versus Internal Control

The rows of Table III represent a second categorization of information about probabilities, namely, external versus internal control of the outcome-generating process. A simple example is the following: in flipping a coin, the outcome is externally determined by various physical factors. In contrast, in throwing a dart at a bull's eye, success depends very much on internal factors such as experience, skills, and effort. Estimating the probability of tails for the coin flip may be done using logical or

TABLE III
Different Kinds of Uncertainty, Depending on Available Information

	Repeated events (frequencies)	*Unique events (scenarios)*
External, environmental control	Relative frequencies of fortuitous or random events, as in casino gambling	System properties, inclinations, strategic plans, policy scenarios, all outside the subject's control
Internal, human control	Relative frequencies of successes and failures, as in repetitive games of skill	Suppositions, ideas, arguments, feelings about the feasibility of one's own plan, task, or ambition

frequentistic data about the coin's likely behavior. Estimating the probability of success in throwing the dart is rather a matter of talent, self-confidence, and determination.

The most relevant contrast in Table III is between repetitive uncertain events under external control (upper left cell), as in pure gambling, and unique uncertain events under internal control (lower right cell), as in the execution of a rare performance task, such as tightrope walking across Niagara Falls.

2.4. When Risks and Benefits Correlate: Risk Homeostasis

In many practical situations, greater benefit may be obtained when one is willing to take higher risk; per activity, benefits and risks often correlate. This was already realized by Coombs and Avrunin in 1977, when they tried to explain so-called single-peaked preference functions. These radically differ from the standard decreasing marginal utility functions adopted by neoclassical economists; see, for example, the positive-utility function in Fig. 2. Examples are the speed of transport, the bravado of mountain climbing, and the difficulty of medical operations.

2.4.1. Optimal Risk Taking

For the example of car driving, Fig. 3 shows positive and negative utility as a function of increasing speed. The upper dashed curve represents the (positive) utility of expected benefits: the faster you drive, the more mobile you are and the earlier you will arrive at your chosen destination; however, there is a decreasing marginal utility of increasing benefits. At the same time, however, the driving costs (e.g., through fuel use and tire wear) and accident risks are increasing disproportionally. The negative utility curve indicates the growing marginal disutility of increasing costs and risks. Thus, as Coombs and Avrunin wrote: "The good satiates and the bad escalates." The solid middle curve in Fig. 3 represents the sum of positive and negative utilities, that is, the overall net expected utility at the various driving speeds. As this sum curve shows, the optimal speed lies somewhere halfway along the speed axis; the subject has a single-peaked preference function for driving speed.

Now let us imagine (not drawn in Fig. 3) that by some external safety measure, for example, the installation of crash barriers along freeways or air bags in cars, the negative utility curve for costs and risks moves slightly
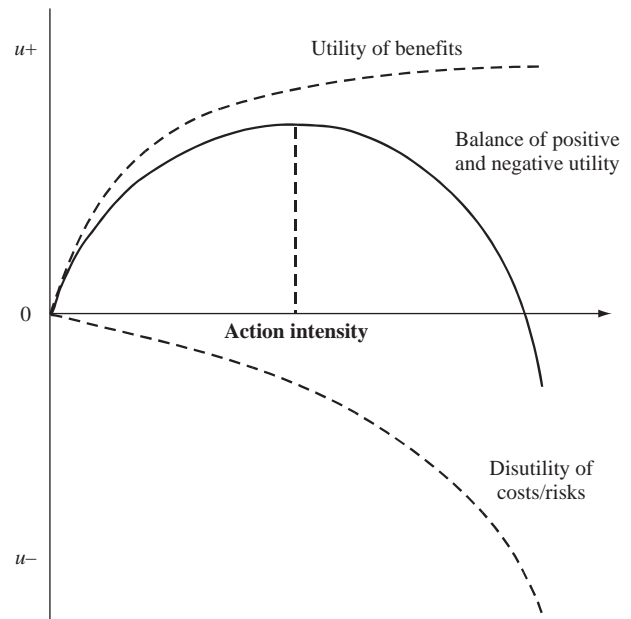

FIGURE 3 Single-peaked utility function (solid curve) for action intensity (e.g., driving speed) as the sum of positive utility of benefits and negative utility of costs and risks, both increasing as action intensifies.

upward. As a result, the summed (solid) expected utility curve would also move slightly upward, so that the optimal car speed is increased (toward the right). This, however, would put the car driver back at roughly the previous negative utility level for costs and risks, thus undermining the intended effects of the external safety measure. In fact, it appears that the subject uses the increased external safety for realizing increased mobility, while the average number of costly incidents and accidents remains the same.

2.4.2. Risk Homeostasis and Safety Regulation

The basic idea expressed by Coombs and Avrunin in 1977, as represented in Fig. 3, has independently been developed in Wilde's risk homeostasis theory. By now, the latter has been well validated, particularly in the domains of road transport, health care, and sports. The great lesson from this work for safety policymakers is that the effects of external safety measures may be canceled by subjects' desire for (even) greater performance benefits.

Risk homeostasis, however, would only apply to activities that leave subjects enough freedom to either intensify their performance or not, as in mountain

climbing, car driving, or performing medical operations. Putting an upper limit to the activity's intensity, for example, by declaring a prohibition on climbing certain mountain slopes or by setting a speed limit on highways, would prevent subjects from using increased safety for achieving greater benefit.

Therefore, increasing the safety of human endeavors seems to require the fulfillment of either one or both of two basic conditions:

1. The extent or intensity of the activity is limited.
2. The performer of the activity enhances his or her intrinsic safety motivation; that is, he or she will perceive the risks as higher than before.

For difficult-to-limit activities such as road transport, challenging sports, and competitive medical care, ultimate safety levels especially depend on the seriousness of perceived risks and the performer's self-control, thereby preventing serious accidents from occurring. On the other hand, when subjects persistently underestimate or belittle the risks, policymakers have no other recourse than to limit the relevant activity.

2.4.3. Risk-Taking Illusions

With reference to Fig. 3, two kinds of illusion may be identified. First, subjects may unrealistically believe that, on further behavior intensification (e.g., faster driving), significantly more benefit may be obtained. Second, subjects may erroneously believe that, on behavior intensification, the costs and risks of the activity will still be manageable. Both the illusion of achieving (still) greater benefit and the illusion of effective risk control may tempt subjects to perform at an higher than optimal level. The end result may be a serious accident involving great loss plus the loss of any benefit originally aspired for. As the saying goes, "It's better to arrive a little late than to end up in hospital."

2.5. Protection Motivation Theory

Used chiefly in health psychology, protection motivation theory (PMT) was developed by Rippetoe and Rogers in 1987 to explain individuals' risk taking and safety policies. PMT is strongly built on cognitive stress theory as developed by Lazarus, Folkman, and others prior to 1984. In 1996, PMT was correspondingly adapted by Gardner and Stern for application to situations of environmental risk. Like cognitive stress theory, PMT comprises a threat appraisal process and a coping appraisal process. See Fig. 4, in which a motivational element has been added to each.

First, threat appraisal rests on the subject's values ("What do I cherish?") and involves perceived severity and vulnerability as well as a judgment of associated benefits; one result is the subject's feelings of fear and/or temptation. Second, coping appraisal pertains to perceived response efficacy ("Could anything be done?"), self-efficacy ("What could I do myself?") and costs and barriers of acting; one result here is the subject's perceived control and his or her motivation to act.

To the extent that the appraisal of coping potential measures up to the appraisal of threat severity, the subject may engage in problem-focused coping aimed at reducing the threat and/or increasing his or her own coping potential. If, however, severity of threat is perceived as overwhelming the subject's coping potential, the subject may (have to) resort to emotion-focused coping, aimed at living with the threat (which may amount to denying it).

Perceived controllability appears to be so important that, psychologically, risk may well be defined as a lack of perceived control in view of significant possible (i.e., non-excludable) negative consequences of an activity or situation. When threat appraisal is low and so is coping appraisal ("Little could be done anyway"), engagement in any other behavior is unlikely. Alternatively, under low threat appraisal and high coping appraisal, subjects would easily take some safety precautions or adopt a safer behavior alternative.

Note that under PMT, subjects have to weigh their coping potential (right side of Fig. 4) up against the appraised threat (left side). This cognitive-affective balancing ("Can I handle the situation?") should somehow accompany the motivational balancing between the utility of associated benefits and the disutility of expected costs and risks, as discussed in Section 2.4 ("Am I willing to take the risks?").

Thus, the essential message of adapted PMT is that risk (and/or stress) management always has two sides. On the one hand, you may work on reducing the threat and the demands of the external environment, which, however, may tempt you via associated benefits. On the other hand, you may work on reducing the vulnerability and enhancing the self-efficacy of the (potential) victims, which, however, may be costly and difficult.

2.6. Emotional Significance of Risk

As marketing specialists very well know, emotions can deeply influence human judgment and decision making.
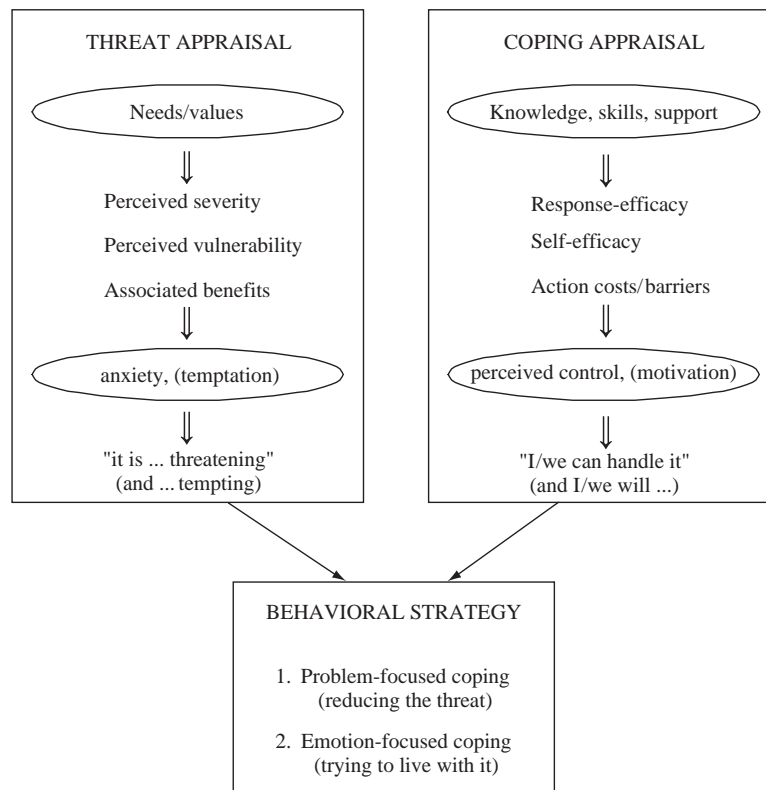
FIGURE 4 Adapted version of protection motivation theory (PMT), with behavioral strategy as a result of the balance between threat appraisal and coping appraisal.

Emotions may be triggered by images of possible decision consequences, which themselves may be based on personal experiences, vivid presentation, proximity in time, and/or cultural conditioning. "Hot" anxiety–sensation thinking is the natural counterpart of "cool" risk–benefit thinking. Just as risks and benefits are balanced cognitively, so must anxiety and sensation be balanced emotionally. This may happen in mutual correspondence, so that cognitions and emotions accompany or even support each other. Not infrequently, however, emotional ("gut") responses may precede and dominate cognitive responses, as may be easily seen in cases of extreme anxiety, fear or time pressure.

2.6.1. Negative Meaning of Risk

Anxiety and fear are prominent emotions in the face of risk. Fear may severely restrict or virtually suffocate deliberate problem solving. Fear of great loss may lead to gambling behavior aimed at trading the greater loss for a smaller one (see Section 2.2). Chronic fear is expressed in phobias, for example, for crowded places, air travel, heights, or certain kinds of animal. Under great anxiety or fear, most people are mentally shortsighted and rather insensitive to rational, economic aspects of risky decision making, especially probability.

The strong priority effect of fear on information processing and (thus) decision quality may block subjects' attention for reassuring information and arguments from more distant authorities. In cases of large-scale involuntary risk, this may cause serious problems of communication between risk-exposed lay subjects and non- (or less-) exposed experts and regulators. Debates about nuclear power generation, liquefied petroleum gas transport, and storage of toxic wastes may provide examples here.

As further elaborated in Section 3.3, subjects' trust in technical experts and responsible authorities is another vital source of emotions. The building of trust not only requires the provision of essential, balanced, and credible information, but also demands that subjects' emotions are recognized and appreciated, as a start toward more deliberative thinking.

2.6.2. Positive Meaning of Risk

Why do people rope-walk across Niagara Falls? Why do they hang-glide from the Eiffel Tower, bungee jump from a high railway bridge into a deep canyon, or start fights with other fans in European soccer stadiums? Why does the public flock to see motor races, air shows, or hazardous circus acts? What explains the increased popularity of so-called survival holiday tours?

As these examples suggest, risk taking can be fun. It can give a kick. Facing risks can challenge our capacities and self-confidence. Successfully completing a risky activity may be a true learning experience. Effectively managing risks may enhance our self-esteem and social status. Without risks, life would be dull and boring. Without a certain willingness to take risks, we would not be able to develop ourselves, to learn new things, and to innovate the physical, social, and economic environment on which future safety and well-being depend.

Also, in line with protection motivation theory (Section 2.5), when coping appraisal is high, that is, when subjects feel capable, skilled, and well-equipped, a given external demand is more easily perceived as a challenge than as a threat, and people's willingness to take risks should be greater.

Even large-scale disasters can have significant positive effects. Afflicted individuals may be forced to get creative and go beyond their everyday limits. Mutual solidarity may be enhanced and social coherence stimulated. After a catastrophe, many people feel that they have endured their ordeal in rewarding togetherness.

3. COMPARATIVE JUDGMENT OF MULTIDIMENSIONAL RISKS

In the past 25 years, a fair amount of research on perceived risk has followed the "psychometric paradigm." The purpose of this inductive approach was to discover the variables and dimensions underlying comparative judgments for a great variety of different risks from industrial activities, transportation, medical treatments, and technologies, as well as household and recreational behaviors. In a much-desired response to limited technical-statistical risk analyses from the engineering disciplines, this work was aimed at answering the following basic questions:

1. If it is not just probability of loss or accident, what (else) is it that causes a particular activity to be perceived as more or less risky and/or acceptable?

2. What might explain the differences in risk perceptions between different groups of people, for example, lay persons and experts, men and women, or people of different political inclination?

3. How meaningful and how feasible are risk comparisons as a basis for risk acceptance and/or the allocation of public funds for relevant scientific research?

3.1. Multidimensionality of Risk

The most powerful conclusion from psychometric research is that risk is a multi-attribute construct whose meaning goes far beyond the classical notions of risk as given in Table II. In the empirical studies underlying this conclusion, various tasks have been presented to respondents. Examples are ratings of risky activities on a set of explicit risk variables, rank ordering of different activities following their perceived riskiness and/or acceptability, and sorting a collection of activities into groups of similar items. The data were analyzed using factor analysis, multidimensional scaling analysis, and/or multiple regression analysis, all capitalizing on patterns of correlation among judgment variables. In addition, theoretical analyses of risky activities were carried out by various other researchers. Based on this varied work, Table IV presents a condensed list of dimensions underlying variations in perceived risk.

The message of Table IV is twofold. First, by rating an activity (insofar as possible) on the 11 risk dimensions, one may understand why subjects consider risks to be as high or low as they seem to believe (thus, the investigator also needs some overall risk criterion measure). Second, the same 11 dimensions hold suggestions of

TABLE IV
Dimensions of Perceived Risk and Corresponding Safety Measures

1. Potential degree of harm or fatality
2. Physical extent of damage (area affected)
3. Social extent of damage (number of people involved)
4. Time course of immediate and delayed effects
5. Probability of undesired consequence
6. Controllability (by self or expert) of situation
7. Experience with, imaginability of consequences
8. Voluntariness of exposure (freedom of choice)
9. Clarity, importance of expected benefits
10. Social distribution of risks and benefits
11. Harmful intentionality

corresponding safety policies to be followed should any dimension be a prominent driver of perceived risk.

3.2. Group Differences in Risk Judgment

The reasons for perceived risk listed in Table IV may be weighted differently by different groups of people. For example, some groups may care less than others about delayed consequences, or they may put unusually high trust in relevant authorities. Major factors underlying intergroup differences in multidimensional risk judgments are the following:

- people's attitudes toward or their interest in the expected benefits involved (cf. dimension 9 in Table IV; see also Section 2.4); a general finding here is that between different activities, perceived risk is lower when associated benefits are larger;
- people's knowledge, capacities, and perceived control as regards the risk (cf. dimension 6 in Table IV; see also Section 2.5); it is in these respects that differences between experts and lay persons may occur, as well as differences between women and men;
- people's world view or general sociopolitical orientation, particularly with respect to authority, technology, and environmental resource use.

3.3. Importance of Trust

As already indicated, perceived controllability (dimension 6 in Table IV) is a crucial risk variable. For the lower level, everyday types of risk following Table I, it is the subject him- or herself who could do much to prevent or control risks and/or to mitigate any accident effects. However, for the higher level, large-scale risks following Table I, many people are dependent on scientific experts, policymakers, and safety inspectors for their information and personal security. Thus, in many cases personal trust in others who are responsible for the matter is a vital component of subjects' feelings of safety.

In 1997, Slovic noted—like others before him—that trust is slowly developed but may be quickly destroyed. Four reasons were listed for this "asymmetry principle":

1. Negative (trust-destroying) events are more visible or noticeable than positive (trust-building) events.
2. When well defined and noticed, negative (trust-destroying) events carry much greater weight than positive events.

3. Sources of bad (trust-destroying) news tend to be seen as more credible than sources of good news.
4. Distrust, once initiated, tends to reinforce and perpetuate distrust.

3.4. Undermining of Self-Confidence

The preceding four reasons were given in connection with technological, social, and environmental risks involving industry, government, and segments of the general public. However, reasons 1, 2, and 4 also pertain to individual risk taking, for example, in traffic, sports, or gambling casinos. Personal losses or accidents stand out and remain cognitively available for a long time, and they thus undermine future self-confidence (reason 1). "Losses loom larger than gains," Kahneman and Tversky noted in 1984, and their greater weight keeps many people from straightforward gain–loss gambling (reason 2). And once in doubt about one's own performance capacities, one may lack sufficient self-confidence for future endeavors (reason 4). Taking this together, we may generally conclude that distrusting oneself takes away the fun or kick, the learning experience, the self-esteem, and the social status of risk taking; cf. Section 2.6.2.

3.5. Meaning and Feasibility of Risk Comparisons

Risk comparisons may be needed

- to decide on the safest way to carry out a given activity (e.g., a medical operation);
- to decide between two or more rather different ways to achieve the same goal (e.g., electric power generation by means of coal, wind, hydro, or nuclear energy);
- to decide on the allocation of investments in various risky activity domains (e.g., road traffic, aviation, food security, sports or medical care) in order to make them safer;
- (and, as we know from above, to find out why and in what respects respondents in psychometric studies find one hazardous activity, situation, or product more risky than another).

Statistical risk comparisons (e.g. between different occupations) have often been made by following a "probability of someone dying" definition, whereby probabilities were estimated on the basis of observed relative frequencies of past deaths. However, the

psychometric research previously discussed has made it clear that in many practical cases fair risk comparisons should be multidimensional. Indeed, multidimensional risk may be formalized (e.g., on the basis of Table IV) in terms of measurements, eventually yielding a bar diagram or a graphic semantic differential representing the risk profile of a particular activity or situation.

Quantitative risk comparisons based on such multidimensional representations are rather difficult to make. They can be completely formalized, and should then follow an explicit evaluation procedure involving attribute (or dimension) selection, activity rating, attribute weighting, and weighted summation of activity ratings across all attributes. Or multi-attribute risk profiles should just be qualitatively presented as such to responsible judges for their intuitive evaluation and rank ordering of activities.

Thus, making risk comparisons is a difficult and often tricky thing to do. In any case, one needs a clear and standardized definition of risk. The safest strategy is to limit risk comparisons to alternative options that fall into the same category (e.g. train versus car; one medical treatment versus another; coal- versus gas-powered electricity generation). Indeed, psychometric risk research itself is based on the somewhat paradoxical assumption that lay respondents can compare multidimensional risks of activities that do not meaningfully fit into a single decision problem.

4. EXTERNAL RISK ACCEPTANCE: RULES, HEURISTICS, AND PROCEDURES

Theoretically, risk acceptance is implied in the choice of a preferred course of (more or less risky) action, whereby expected benefits constitute the primary motivating factor (see Section 2.1). In many cases, however, risk-exposed people do not directly enjoy the benefits of the relevant activity. Under such involuntary (or nonbeneficial) risk exposure, explicit rules and procedures may be needed to determine the acceptability of a given risk situation.

4.1. Risk–Benefit Trade-Offs

Whether we deal with a static gambling situation (as in Section 2.1) or with a dynamic control situation (Section 2.4), risk taking often requires a weighing of risks against benefits. The larger the benefit involved, the higher the risk accepted. So, if you see someone voluntarily take a high risk, always ask yourself which compensating (or seductive) benefit the subject expects to achieve. Of course, several other factors may play a significant role as well, for example, a possible unrealistically optimistic underestimation of the probability of loss ("It won't happen to me"), time-discounting of long-term risk ("Who cares about later?"), and an exaggerated need for the benefit involved, for example the "kick" from a hazardous sporting event. For society at large, important economic benefits often justify the toleration of appreciable risks.

4.2. Risk-Limiting Standards

For the sake of cognitive convenience, in many practical situations risk–benefit trade-offs are "frozen" into simple risk acceptance standards. Examples are vehicle speed limits on highways, the degree of attentiveness in descending a staircase, or the limitation of preservatives in food products. In such cases, without giving up the intended benefits, one is willing (or forced) to adapt the activity, at some explicit or implicit cost, in order to make it safer. Acceptance standards often have a third-party protection function, too, as illustrated by speed limits on highways.

4.3. Risk as Low as Reasonably Achievable

Some consequences of a risky activity may sooner or later be fatal. Thus, one should naturally be willing to reduce such risks to a level that is "as low as reasonably achievable." This ALARA principle, however, may lead to a kind of phobia. For example, some may refuse to eat shellfish, some would never take an airplane trip, and some, knowing they are a carrier of a genetic disease, would decide to never have children. The ALARA principle means that personal safety strategies, although primarily benefit-driven and perhaps somehow standardized in terms of risk limits, may further involve efforts to reduce the probability of some sad loss as much as possible, even if the subject considers the situation to be safe enough otherwise. The result, however, may be that an activity becomes very costly or no longer feasible.

4.4. Justification, Risk Limits, and ALARA

Risk–benefit weighing, the testing of risk acceptance standards and subsequently trying to get risks ALARA, may constitute an effective combined strategy for risk management. For example, for involuntary risks to radiological workers (e.g., in hospitals and nuclear energy laboratories), the International Commission on Radiological Protection has recommended the use of such a joint set of safety principles. This strategy comes close to an empirical conclusion drawn by Hendrickx in 1991, that in a dynamic control task, risk acceptance consists of a deliberate choice of a safe-enough course of action followed by an allocation of effort to keep the risks under sufficient control. A triple heuristic conclusion on risk acceptance was also drawn on empirical (psychometric) grounds by Vlek and Stallen in 1981: for most lay people, risk acceptance hinges on: (1) the importance of expected benefits, (2) limitation of the maximum possible accident, and (3) the perceived controllability of the activity as a whole.

4.5. Societal Risk Acceptance: Importance of Procedures

Involuntary exposure to industrial or environmental risks easily raises public concerns, protests, and feelings of being treated unfairly. To the extent that expected benefits are the main driving force of risk acceptance, we may expect involuntary and mostly nonbeneficial exposure to cause much trouble for responsible decision makers. "NIMBY," or "not in my back yard" is one protest often heard regarding the question of where to build hazardous facilities such as municipal waste incinerators, electric power stations, and airport expansions.

The reactions of involuntarily risk-exposed persons toward responsible authorities have been widely studied, and various suggestions have been formulated to facilitate the processes of decision making and eventual risk acceptance. Table V lists six different diagnoses of public protests together with recommended policy approaches.

Specialists believe that in most practical cases, several diagnoses may be valid at the same time. Hence, they recommend that administrative procedures for planning, decision making, and undertaking of a hazardous activity involving involuntary risk should follow a carefully designed procedure incorporating all or most of the therapeutic ingredients mentioned in Table V.

TABLE V

Six Possible Diagnoses and Corresponding Therapeutic Approaches Toward "NIMBY" ("Not in My Backyard") Attitudes

Diagnosis	Policy approach	Labelling name
1. Ignorance	Communication	Technocratic approach
2. Unfairness	Compensation	Market approach
3. Exclusion	Participation	Participation approach
4. Social inequity	Redistribution	Social justice approach
5. Policy criticism	Strategic alternative	Transformation approach
6. Pure local egoism	Government coercion	Regulatory approach

5. COLLECTIVE RISKS IN COMMONS DILEMMAS

Involuntary exposure often occurs in connection with environmental risks at the higher scale levels following Table I. In quite a few cases, collective risks are generated in a process of accumulating small contributions from numerous individual actors.

5.1. Definition of a Commons Dilemma

A commons dilemma is a situation of conflict between an aggregate collective interest and numerous individual interests; see Hardin's inspiring 1968 analysis of the disastrous consequences of common pasture land ("the commons") being overgrazed by ever more cattle owned by individual farmers, each trying to improve their individual income. A present-day example is the urban air pollution resulting from accumulation of exhaust fumes from many individual cars.

5.2. Macro and Micro Level Perspectives

At the macro level of society, a commons dilemma presents itself as a permanent contrast between a collective risk, such as air pollution, and a large collection of individual benefits, such as the benefits of car driving, for example. Minimization of risk and maximization of benefits are incompatible social goals,

between which a trade-off must be made. In contrast, individual actors at the micro level of society are naturally focused on their own benefits "here and now". Individuals may not recognize a commons dilemma as such

- because they are unaware of any collective damage;
- because they do not appreciate their own responsibility for collective problems;
- because they do not feel that the long-term collective risk is serious enough in relation to the numerous short-term benefits; and/or
- because they know there is a collective risk, but feel that little can be done about it.

In particular, the combination of short-sightedness, personal benefit, and perceived lack of control over the situation may readily lead to a denial or playing down of the risk and thus to a refusal to participate in risk control efforts.

5.3. Key Issues for Research and Policymaking

Understanding commons dilemmas and managing the collective (environmental and social) risks generated by individual activities revolve around problem diagnosis, policy decision-making, practical intervention, and evaluation of policy effectiveness. These four categories can be expanded into 12 focal issues, as listed in Table VI. The 12 issues listed in Table VI imply rather different tasks to be performed by experts from different backgrounds, not just psychologists. Also, the 12 activities would have to be carried out in reasonably good order, so that subjects of intervention strategies understand why they are asked to accept or cooperate, and to what future improvements this might lead.

5.4. Structural and Individual Factors of Social Cooperation and Defection

A multitude of experiments and field studies have clarified the main factors governing people's tendency to either cooperate or defect (i.e., behave selfishly) when faced with a commons dilemma. Table VII reviews the structural and individual psychological factors governing cooperation.

In connection with this, seven general strategies may be used for the promotion of cooperation in managing

TABLE VI

Twelve Key Issues for Research and Policymaking on Commons Dilemmas

I. Problem diagnosis
  A. Analysis and assessment of collective risk, annoyance, and stress
  B. Analysis and understanding of sociobehavioral factors and processes underlying risk generation
  C. Assessing problem awareness, risk appraisal, and actors' individual values and benefits

II. Policy decision making
  A. Weighing of collective risk against total individual benefits ("need for change?")
  B. If "risk unacceptable": setting objectives for reduction of environmental and/or social risks
  C. Translation of risk reduction objectives into individual behavioral goals

III. Practical intervention
  A. Focusing on individual target groups and considering essential conditions for policy effectiveness
  B. Specifying feasible behavioral alternatives and selecting effective policy instruments
  C. Targeting group-oriented applications of strategic programs of behavioral change

IV. Evaluation of effectiveness
  A. Designing a monitoring and evaluation program to determine policy effectiveness
  B. Systematic, comprehensive evaluation of observable effects and side effects
  C. Intermittent and *post hoc* policy feedback and possible revision of policies

collective risks. These are briefly indicated in Table VIII, along with one default strategy.

Policymakers regularly use separate strategies in a piecemeal fashion. However, they rarely apply a comprehensive approach to managing collective risks in commons dilemmas, involving problem diagnosis, policy decision making, and practical intervention following various related strategies.

5.5. Precautionary Principle

Large-scale commons dilemmas (e.g., motorized transport or genetic engineering) may yield great environmental risks that may be hard to characterize. Not only is there a variety of threatening consequences whose time of occurrence is rather uncertain, but also there is great ambiguity about the probabilities of catastrophic

TABLE VII

Factors Facilitating Cooperation to Reduce Collective Risks in Commons Dilemmas

Structural factors	Individual psychological factors
• Smaller group	• Environmental problem awareness, insight into causal processes
• Stronger group identity	• Feeling of responsibility, moral obligation
• Higher visibility of others' behavior	• Perceived (relative) appeal of cooperation
• Availability of cooperative alternatives	• Self-efficacy in making cooperative contributions
• Lower cost of cooperative choice	• Trust in others' cooperation (low social uncertainty)
• Higher cost of defective choice	• Ability to resist bad social examples ("rotten apples")
• Legal obligation to cooperate	• Perceived environmental effectiveness of cooperation (outcome efficacy)
	• Altruistic and cooperative social value orientation

events. What people do realize is that possible catastrophes are caused by humans and so should eventually be controlled by humans.

For environmental policymakers, a reasonable suspicion of a serious threat may offer sufficient grounds for applying the precautionary principle. This means that various risk avoidance, prevention, and/or reduction actions are implemented, aimed at reducing the threat without there being sufficient scientific evidence to conduct a justifiable risk–benefit analysis. Because this is a controversial practice, psychologists might contribute to the debate by investigating the extent to which everyday safety decisions involve precautionary

TABLE VIII

General Strategies for Behavioral Change in Managing Commons Resource Dilemmas

- Provision of physical alternatives, (re)arrangements
- Regulation and enforcement
- Financial-economic incentives
- Information, education, communication
- Social modelling and support
- Structural, physical, and/or social change
- Changing values and morality

measures either preceding or falling outside well-informed rational decision making.

6. CONCLUSION

Both small- and large-scale levels of risk offer their own problems of analysis, judgment, decision, and control. Many risks are difficult to reliably quantify, due to lack of experience and information and because of the dynamic nature of the relevant activity. Thus, expert judgment, perhaps that of the subject assessing the risk, becomes an important factor, and total process control becomes more important than single gambling types of acceptance decision. Therefore, psychologists may define risk as a lack of perceived control (or of coping potential) in view of significant possible negative consequences.

In many cases, an activity or situation is considered to be risky for a variety of reasons (see Table IV). Multidimensionality implies that these activities or situations could be made (to appear) more safe in a corresponding variety of ways. Thus, classical definitions of risk as a function of probability and size of loss have limited practical significance. Multidimensionality of risk also implies that one should be very careful in making risk comparisons across different domains: comparisons variable(s) should be similar and well defined.

As regards large-scale social and environmental risks, sophisticated strategies are possible to limit such risks in view of the numerous social and economic benefits involved. The commons dilemma is a useful model that allows comprehensive problem diagnosis, policy decision making, and well-tuned practical interventions. For highly ambiguous environmental threats, the precautionary principle may be invoked, but this needs psychological investigation to support its practical applicability.

Finally, for future studies on risk, psychologists should familiarize themselves more deeply with specific domains such as industry, transportation, health care, and environmental protection to make their work less abstract and give it more practicality. This would also enable psychologists to work together with experts from other disciplines, whose concepts and tools may be instructive but will likely also need psychological enrichment.

See Also the Following Articles

Environmental Design and Planning, Public Participation in ■ Safety and Risks: Errors and Accidents in Different Occupations ■ Trust ■ Work Safety

Further Reading

Gardner, G. T., & Stern, P. C. (1996). *Environmental problems and human behavior*. Boston, MA: Allyn and Bacon.

Gärling, T., Biel, A., & Gustafson, M. (2002). The new environmental psychology: the human interdependence paradigm. In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 85–94). New York: Wiley.

Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, *39*, 341–350.

Klinke, A., & Renn, O. (2002). A new approach to risk evaluation and management: risk-based, precaution-based and discourse-based. *Risk Analysis*, *22*, 1071–1094.

Lazarus, R., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Liebrand, W. B. G., & Messick, D. M. (Eds.) (1996). *Frontiers in social dilemmas research*. New York: Springer Verlag.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Rippetoe, P., & Rogers, R. (1987). Effects of components of protection-motivation theory on adaptive and maladaptive coping with a health threat. *Journal of Personality and Social Psychology*, *52*, 596–604.

Slovic, P. (1997). Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. In M. H. Bazerman, D. M. Messick, A. E. TenBrunsel, & K. A. Wade-Benzioni (Eds.), *Environment, ethics and behavior. The psychology of environmental valuation and degradation*. San Francisco, CA: The New Lexington Press.

Vlek, C., & Keren, G. B. (1992). Behavioral decision theory and environmental risk management: assessment and resolution of four 'survival' dilemmas. *Acta Psychologica*, *80*, 249–278.

Wilde, G. J. S. (1994). *Target risk*. Toronto, Canada: PDE Publications.

Yates, J. F. (1992). *Risk-taking behavior*. Chichester/New York: Wiley.

Ethics and Social Responsibility

Pieter J. D. Drenth

Vrije University, Amsterdam, The Netherlands

GLOSSARY

"no-go" decisions A term meaning that the research in question is wholly unacceptable with respect to specific research questions on the ground of social and ethical objections.

"slow-go" decisions A term that may include cases in which there is insufficient knowledge to properly appreciate and control the social or moral consequences of the research with respect to specific research questions on the ground of social and ethical objections; in such instances, results should be temporarily suspended until the ethical implications have been subjected to public discussion and consensus or at least a reasonable majority agreement has been reached.

Wissenschaf als Inhalt A morally neutral term referring to science as a product (i.e., the accumulated body of knowledge) solely subject to methodological–scientific norms.

Wissenschaft als Tätigkeit A term referring to science as the process of knowledge accumulation subject to ethical and political norms.

For quite some time, academic psychologists have incurred risks of accusations of "superbia" by developing an "ivory tower" attitude in which any accountability for human and social effects of their research is repudiated: "Science is about how things are, not about they should be." During the past few decades, however, the question of science and values has become a major subject of discussion again—certainly within psychology. Attacks on the autonomy and sovereignty of scientific psychology have come from various sources. First, there was the antiestablishment movement of the 1970s in which the political–scientific reflections of authors such as Marcuse, Adorno, Habermas, and Holzkamp became quite popular and were willingly embraced by student activists and (mostly junior) staff. Their criticism certainly contributed to the dismantling of the misconception that freedom of science would imply a negation of societal responsibility. A second attack resulted from the economic recession in many Western countries during the 1980s that led to a severe reduction of public and private funding for research. Acknowledged (and published) research output became an important condition for the survival of research personnel. This challenged another misconception of scientific freedom, namely, the license to freely think, meditate, and explore without caring much about output or quality. A third attack on the free and autonomous academic psychology came from the cry for "utilization" of scientific research at the end of the 1980s and during the 1990s that was heard in both governmental and industrial circles: Scientific goals should be seen as subordinate to those of economic and technological development, and criteria such as utility, applicability, and economic relevance should be given priority over pure scientific standards. Industrial, societal, and/

or technological relevance became an important consideration for funding research. It is ironic how much the neo-Marxist principles of the 1970s coincide with the industrial utility touchstones of the 1990s. Fourth, during recent years, the question of the autonomous and value-free character of science versus the relevant and value-bound character of science has received ample attention. A more detailed discussion of this issue is presented in this article.

1. SCIENCE: VALUE FREE OR VALUE BOUND?

The question phrased in the heading of this section has become an important issue during the past few decades, although opinions still diverge. On the one hand, many defend that scientific knowledge is value free and has no moral connotations. Science tells us how the world is, whether we like the story or not. Basic research is driven by scientific curiosity and not by the hope that it will be put to practical use. Ethical and moral issues can arise only when science is applied and is expected to produce usable objects. But then it has become technology and is not science anymore. Technological objects or processes can be used for better or for worse. Science produces insights, ideas, and pieces of knowledge that are in themselves neutral and can be corrupted only if they are mixed with political, social, economic, or other nonscientific aims.

On the other hand, there is a different view that does not accept the premise that science should be concerned only with producing reliable knowledge and, consequently, should be value free. The following arguments for the value-bound character of science can be brought forward.

• It is a basic obligation of all scientists and scholars to reflect on the paradigmatic presumptions and sociohistorical entrenchment of their scientific activities. This reflection is, in itself, a metascientific and value-embedded phenomenon. Scientists must realize that their conceptualizations and models are always abstractions of reality and that only an approximation—or "reconstruction"—of reality can be achieved.

• Recently, the distinction between basic science and applied science has become less clear-cut. There is a good deal of overlap between the two spheres, and it is increasingly difficult to identify parts of science that do not affect, or are not affected by, technology. In this light, it will become clear that reserving "value-free

autonomy" for science and "value-bound heteronomy" for applied science and technology is no longer tenable.

• Scientists, and especially behavioral scientists, deal with a social, political, or psychological reality that is continuously affected and changed by the very scientific findings they produce. Health, safety, communication, privacy, mobility, welfare, economic development, and many other valuable goals of humankind are influenced radically by the advances of modern science. At the same time, many ethical or sociopolitical problems result directly from these same advances. Scientists should be aware of this and should anticipate the changes they bring about and the problems these changes raise.

• Even if scientists refrain from actually making political or ethical choices and restrict themselves to presenting probabilities and risks coupled with certain options, their reasoning is not value free. Risks do involve values and normative choices that scientists must face. Questions such as the following arise. Risks for whom? How far does the "right to know" go? What is the balance between self-determination of the individual and the interests of larger groups or the society as a whole? At what level of certainty does the scientist have to warn, especially in case of irreversible developments?

• Scientists cannot avoid the metascientific question of whether or not it is worth knowing what they pursue. They must justify—not only for themselves but also publicly if taxpayers' or sponsors' money is involved—why scientific issues must be addressed. This justification implies, in essence, nonscientific choices and decisions.

These arguments warrant the following conclusion in scientific research. On the one hand, objectivity must be maintained against the pressure from ideological movements, industrial lobbies, governmental strain, and/or political pressure groups. On the other hand, it has become ever more difficult to separate the functions of producing knowledge and making value-bound choices in extending research findings to the public or to the society at large. Here it may be helpful to distinguish between science as a product of knowledge (i.e., the accumulated body of knowledge) and science as a process of knowledge accumulation. The former, which was referred to as *Wissenschaf als Inhalt* by the Swiss philosopher Bochenski at the 1990 Engelberg Forum, is morally neutral and subject solely to methodological–scientific norms. The latter, which Bochenski described as *Wissenschaft als Tätigkeit,* is

subject to ethical and political norms that stipulate the choice of hypotheses to be researched, the manner in which data are gathered and experiments are conducted, and the question of what is ultimately done with research data.

Thus, research is embedded in the context of values, interests, and political objectives. Rather than denying this or retreating to the safety of the ivory tower, scientists would do well to realize this and take the appropriate responsibility seriously.

2. ETHICAL AND SOCIAL RESPONSIBILITY

This article does not discuss the ethical and social responsibilities of the professional psychologist in everyday practice. It is clear that in the application of psychology, be it in a clinical, pedagogical or social/organizational setting, the professional worker comes across many moral problems and dilemmas. Most national psychological associations have dealt with these issues, and many have developed codes of ethics for the professional application of psychology in various settings. A discussion of these codes falls outside the scope of this article.

But it has become clear that if one looks at the ethical and social responsibility of the psychologist as a scientist, his or her liabilities do not stop at the gate of the laboratory. This responsibility reaches further than the limits of purely scientific norms. It concerns the choice of the subject to be researched (or not), the proper conduct of experiments and surveys with due respect for human participants and/or animal subjects, and accountability for the application and use of the scientific results. More specifically, the following prescripts are obtained. One can defended that the psychologist as a scientist should do the following:

- Promote and take an active part in the national and international discourse on ethical and social conditions and consequences of scientific research in psychology and its implementations.
- Be on the alert for selective and biased interpretation of the research findings by opponents or advocates of this research and deprecate such bias whenever discovered.
- Be observant of possible improper use or abuse of the research findings, warn against such possible misuse, and take corrective actions wherever possible.

- Provide careful, complete, and understandable information about the research results, not only in scientific journals but also in professional magazines and popular media (if possible). The latter should be considered important particularly if the research tends to have a substantial social, political, or economic impact.
- Promote the development and the acceptance of a set of rules and regulations for ethical conduct for professional practice in psychology, and train and educate students and young professionals to obey these rules and to develop conscientious attitudes and integrity in applying research results in their future professional work.
- Advocate and develop conditions that inhibit improper applications of scientific products (e.g., tests, instruments, treatments, intervention methods), including restrictive sales, extension of copyrights, and penal sanctions.

In other words, the responsibility of the psychologist as a scientist would have to be translated into very practical efforts and activities and specific rules of conduct.

3. "NO-GO" OR "SLOW-GO" DECISIONS?

An interesting question is whether scientific curiosity should be uninhibited and scientific research should be unrestrained, or whether social or ethical constraints should be imposed on science in some cases—in other words, whether there could be a need for "no-go" or "slow-go" decisions with respect to specific research questions on the ground of social and ethical objections. No-go means that the research in question is wholly unacceptable. At this time, there are cases in which such a constraint would find wide support, including attempts to make human–ape hybrids, applications of germline genetic manipulation to enhance intelligence, and human cloning for procreation. Slow-go would apply in instances where results should be temporarily suspended until the ethical implications have been subjected to public discussion and consensus or at least a reasonable majority agreement has been reached. This area might include cases in which there currently is insufficient knowledge to properly appreciate, and thus control, the social or moral consequences of one's research.

It should be noted first that any discussion on constraints to be imposed on research is fraught with

danger. Throughout history, there have been too many examples of scientists who were repressed and whose research was stifled because the research outcomes were not consonant with the ruling ideologists, did not curry favor with powerful authorities, or were opposed to the interests of influential movements and action groups. Even if some of these movements have reputable objectives, such as peace and détente, equal rights for women, environmental protection, and non-discrimination, infringement of the right to investigate on the basis of political unacceptability of possible outcomes to external groups is highly precarious.

It would be equally wrong to refrain from doing research in a given area because it might possibly be abused or applied irresponsibly. This would virtually mean the end of all scientific research given that no scientific result is secured against willful abuse.

Therefore, it is not easy to identify incontestable ethical constraints to scientific research. Such constraints should refer to basic and peremptory values that would be imperative for scientists and responsible citizens alike. With respect to this issue, the European Federation of National Academies of Sciences and Humanities has discussed and accepted the following principles that could also be endorsed by researchers in psychology:

1. Research is not justifiable if unacceptable damage is inflicted on the object of research or its wider environment (e.g., social environment, society, nature) before, during, or after the experiment or the gathering of data. This applies to all objects of research, whether they be people, animals, nature, or culture.

2. Research is not justifiable in cases where the nature or the consequences of the research are in conflict with basic human values. These values include, in any case, (a) respect for human dignity, which guarantees the autonomy and freedom of choice of all individuals, informed consent prior to participation in research, and the rejection of every intent to commercialize the human body; (b) solidarity with humankind, which guarantees regard and acceptance of fellow humans on the basis of equality; and (c) solidarity with posterity, which embodies the broader responsibility for sustained development of a planet that is to be left to future generations.

See Also the Following Articles

Industrial/Organizational Psychology, Overview ■ Occupational Psychology, Overview

Further Reading

Berthoud, G., & Sitter-Liver, B. (1996). *The responsible scholar*. Canton, MA: Watson Publishing International.

Drenth, P. J. D., Fenstad, J. E., & Schiereck, J. D. (1999). *Science and scientists between freedom and responsibility*. Luxembourg: Office for Official Publications of the European Community.

Heller, F. A. (1986). *The use and abuse of social science*. London: Sage.

Kimmel, A. J. (1988). *Ethics and values in applied social research*. Newbury Park, CA: Sage.

Nagel, T. (1979). *Mortal questions*. Cambridge, UK: Cambridge University Press.

Shea, W. R., & Sitter-Liver, B. (1989). *Scientists and their responsibility*. Canton, MA: Watson Publishing International.

Exceptional Students

Jessica Singer-Dudek

St. John's University, Jamaica, New York, USA

GLOSSARY

applied behavior analysis A scientific method for educating all students; its basic premise is that all behavior is controlled by its consequences, and so all behavior can be changed.

continuum of services The range of special education placement options, ranging from the least restrictive environment to the most restrictive environment.

functional behavior assessment A critical method for determining the function or purpose of behavior; it involves determining not only the consequences of behavior but also the antecedents of or precursors to behavior.

high-incidence disabilities Specific learning disabilities, speech and language impairments, mental retardation, and emotional disturbance.

inclusion When an exceptional student is a member of a regular class and receives instruction in the regular education setting for the majority of the school day; it may also refer to the philosophy of educating all students in a nonsegregated environment.

individualized education plan (IEP) A plan whereby a contract is devised among the school district, educational placement (i.e., the school), and a parent(s) outlining a student's educational plan for the school year; it includes the type of educational placement, services to be provided, progress from the past school year, recent test scores and evaluation results, and goals and objectives for the school year.

low-incidence disabilities Nine disabilities identified under the Individuals with Disabilities Education Act (IDEA): other health impairments, multiple disabilities, autism, orthopedic impairments, hearing impairments, developmental delays, visual impairments, traumatic brain injury, and deafness/blindness.

mainstreaming When an exceptional student is a member of a special class but attends a regular class for certain periods of the day; a student with special needs can be mainstreamed for academic or nonacademic subjects.

reinforcement The principle of behavior stating that a desirable consequence delivered immediately after a behavior will serve to increase the likelihood that the behavior will occur again in the future.

related services Supplemental services that may include speech and language therapy, physical therapy, occupational therapy, nursing services, counseling, hearing services, and/or visual services.

self-contained classroom A special class that consists solely of students with disabilities; it has a smaller teacher-to-student ratio and may include teaching assistants or paraprofessionals.

self-fulfilling prophecy When low expectations are set and students only learn to meet these lower standards.

self-injurious behavior Repetitive or nonrepetitive behavior that results in bodily harm to the individual, either immediately or over time (e.g., scratching, hair pulling, face slapping, scab picking, head banging).

stereotypy Repetitive behaviors that appear to have no function other than the sensory enjoyment of the behaviors themselves; stereotypic behaviors are often emitted by students who are diagnosed with autism or mental retardation.

Exceptional students comprise nearly 13% of all American school-aged children. Approximately 10% of students in the United States are diagnosed with disabilities, whereas 2 to 3% of the population is characterized as gifted or talented. The education of students with disabilities in the United States is governed for the most part by Public Law 94-142, also known as the Individuals with Disabilities Education Act (IDEA). Although exceptional students can be separated into 13 disability categories under IDEA, each student is unique. The learning, social, and behavioral characteristics of exceptional students, including gifted and talented students, differ for each individual student. The individualized educational needs of exceptional students must often be determined using alternative types of assessments. According to these assessments, the most appropriate placement, services, and instructional modifications are put into place for each student. Exceptional students can be educated in a range of educational settings, including the regular classroom. Under IDEA, students with disabilities receive an education according to an individualized education plan (IEP) designed around their unique needs and goals. Special educational plans may also be arranged for gifted or talented students. Because of the careful identification of goals and individualization of instruction, collaboration between professionals and families is critical for the successful education of exceptional students.

1. WHO ARE EXCEPTIONAL STUDENTS?

Exceptional students comprise the population of students with disabilities who require specialized educational services as well as students who are characterized as gifted or talented. Approximately 10% of students in the United States are diagnosed with disabilities, whereas approximately 2 to 3% of the population is considered to be gifted or talented.

The term "exceptional" has often been used to describe unusual, unique, or outstanding qualities of people or objects. Consider the following phrases: "His artwork is exceptional" and "She is exceptionally bright."

In this case, the term "exceptional" refers to students who learn and develop differently from most others or students who have exceptional learning styles, exceptional talents, or exceptional behaviors. Exceptional students are those who fall outside of the normal range of development. Some disabilities are identified at birth or soon after birth, whereas others go undetected until the students enter school, where learning problems first become apparent. In the case of gifted or talented students, their abilities may also go undetected until they reach school. Some exceptional students have learning problems, some have behavioral problems, some have communication difficulties, and some have social deficits. Many students with exceptionalities have two or more of these problems, although they may also excel in certain areas. Students who are gifted or talented may also display behavioral or social difficulties. For exceptional students, the most important distinctions are made in the area of educational needs.

2. THE INDIVIDUALS WITH DISABILITIES EDUCATION ACT

The education of exceptional students in the United States is governed by a federal law known as the Individuals with Disabilities Education Act (IDEA), which was first enacted in 1975 and was then known as the Education of All Handicapped Children Act. Prior to 1975, U.S. schools were not mandated to educate students with disabilities. Very often, students with disabilities were deemed to be uneducable and were barred from entering schools. Now, their educational rights are protected, and dictated, by IDEA. Students are protected under IDEA from 3 years of age until the end of the school years in which they reach their 21st birthdays unless the students graduate earlier. The education of students who are exceptionally gifted or talented is governed by individual states. However, not all states have legislation pertaining to the education of gifted or talented students.

According to the U.S. Department of Education, a total of 5,762,935 students with disabilities ages 6 to 21 years were served under IDEA during the 2000–2001 school year. IDEA recognizes 13 disability categories. The number of students in each disability category served under IDEA during the 2000–2001 school year is listed in Table I. Each category of exceptionality has its own set of criteria for diagnosis. Students must meet all criteria to be diagnosed with a particular disability.

Disability category	Number of students served
Specific learning disabilities	2,879,445
Speech and language impairments	1,092,105
Mental retardation	611,878
Emotional disturbance	472,932
Other health impairments	291,474
Multiple disabilities	121,954
Autism	78,717
Orthopedic impairments	73,011
Hearing impairments	70,662
Developmental delay (ages 3–9 years)	28,683
Visual impairments	25,927
Traumatic brain injury	14,829
Deafness/Blindness	1,318
All disabilities	5,762,935

Source. From U.S. Department of Education, Office of Special Programs, Data Analysis System. (2002). *To assure the free appropriate public education of all children with disabilities: Twenty-fourth annual report to Congress on the implementation of the Individuals with Disabilities Education Act.* Washington, DC: U.S. Department of Education.

Sometimes, students are impaired in more than one area, and these students may receive a diagnosis of multiply handicapped. The first four disability categories—specific learning disabilities, speech and language impairments, mental retardation, and emotional disturbance—are considered to be high-incidence disabilities because they affect the majority of students with disabilities. Specific learning disabilities affect nearly half of all students with disabilities. The other disability categories are considered to be low-incidence disabilities.

2.1. The Six Principles of IDEA

Exceptional students often require specialized education to meet their unique needs or to accommodate and build on their special talents. The majority of students with disabilities and students who are gifted or talented require special education. IDEA ensures a free and appropriate public education (FAPE) to all students with disabilities. Under IDEA, there are six important principles that dictate how students with disabilities are diagnosed, evaluated, placed in educational settings, and educated. These principles are discussed in the following subsections.

2.1.1. Zero Reject

The zero-reject principle ensures that no student can be denied an education due to the presence or severity of a disability. A student with a severe learning difficulty or severe behavior problem may not be denied services or expelled from school due to his or her disabilities. Under this principle, a student who is terminally ill or has a contagious disease may not be denied a FAPE either.

2.1.2. Nondiscriminatory Evaluation

The second principle of IDEA ensures nondiscriminatory evaluation of any student who has, or is suspected of having, a disability. Several types of assessment are conducted by an interdisciplinary team of nonbiased members. No evaluations may be conducted without parental consent. All decisions pertaining to a disability diagnosis and to educational evaluation are made by a committee that includes the parent(s) or guardian(s) of the child. The committee seeks to answer three important questions. Does the student have a disability? If so, does the student require special education services? If so, what educational placement and which services will best serve the student's needs?

2.1.3. Appropriate and Beneficial Education

The third principle of IDEA ensures an appropriate and beneficial education. This means that the student must be placed in an educational setting that best suits his or her needs. The student must be making educational progress for the placement to be deemed beneficial. An individualized education plan (IEP) is devised for each student, outlining the progress, placement, services, and goals to be addressed for the student. Appropriate placement means that the student is placed in the least restrictive environment (LRE).

2.1.4. Least Restrictive Environment

IDEA requires that the student be educated with his or her nondisabled peers to the maximum extent possible. The LRE is the educational setting that is closest to the regular educational setting that can still meet the student's individual needs. Only if a student's needs are too great to be met in the regular education environment should a more restrictive placement be considered. The range of educational placement options is known as the continuum of services.

2.1.5. Procedural Due Process

The fifth principle of IDEA ensures procedural due process. This means that there are safeguards in place in case the parents and school district disagree about what is best for the student. Impartial officials preside over due process hearings and make decisions according to what is best for the student. The principle of procedural due process also dictates how parents are informed of all decisions regarding their child, including the child's educational progress.

2.1.6. Parent and Student Participation

Parent and student participation is encouraged and mandated under the sixth principle of IDEA. This means that no decisions shall be made regarding the education of an exceptional student without parental consent and input. Student input is also encouraged whenever the student is capable of contributing. The student should have the opportunity to make decisions about his or her own future.

2.2. The Continuum of Services

Under IDEA, students with disabilities must be educated in the LRE, which is the educational environment that is closest to the general education classroom and that still meets the students' needs. There is a range of educational placements for students with disabilities, known as the continuum of services. Figure 1 illustrates the continuum of services.

At the top of the spectrum is the general education classroom. The student may receive supplemental services, known as related services, while in this placement (e.g., speech/language therapy, physical therapy), but
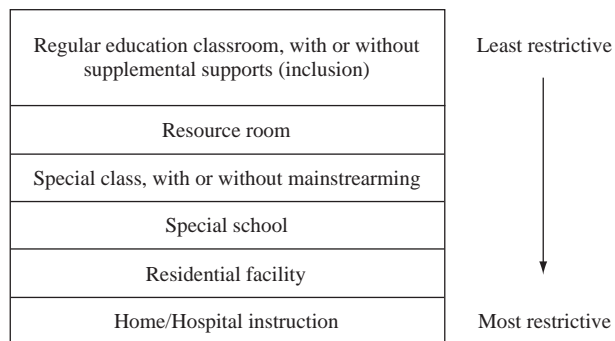
Regular education classroom, with or without supplemental supports (inclusion)	Least restrictive
Resource room	
Special class, with or without mainstrearming	
Special school	
Residential facility	
Home/Hospital instruction	Most restrictive

FIGURE 1   The continuum of services from the least restrictive environment to the most restrictive environment.

the student is a member of the regular class. When a student with a disability is a member of a regular class and spends the majority of the school day in that classroom, this is known as inclusion. An inclusion setting is an appropriate and beneficial placement only if the student is making educational progress. The social benefits of the inclusion setting are also taken into account and weigh heavily when determining whether an inclusion setting is the best placement option.

The next option is a resource room. Here, the student is a member of the regular class but spends up to two periods each day in a special class taught by a special education teacher. Usually, particular goals are worked on in the resource room (e.g., math, reading, study skills). The next placement option on the continuum of services is a special class or self-contained setting. Here, the student is placed in a class consisting solely of students with special needs. Special classes have smaller student-to-teacher ratios and usually include one or two teaching assistants or paraprofessionals. There are a variety of student-to-teacher-to-teaching assistant ratio options available. Sometimes, students who are placed in self-contained classes actually excel in one or more subject areas. For example, a student may require specialized instruction for reading and all content-area subjects but be on grade level for math. This child may attend a regular math class during the school day. This is known as mainstreaming. In keeping with IDEA, the student is spending as much time as possible with nondisabled peers.

The next placement option is a special school. Here, the student attends a school composed solely of students with disabilities. This restrictive setting ensures that programs and accommodations are available consistently. Beyond a special day school is the next placement option on the continuum of services, that is, a residential facility. The student attending a residential facility requires 24-hour continuation of instructional programs and care. The final placement on the continuum of services is home or hospital instruction. Here, instruction must be provided by a special education teacher for at least some portion of the school day.

3. LABELING AND CLASSIFICATION OF EXCEPTIONAL STUDENTS

There are several advantages and disadvantages to labeling students with disabilities. The first advantage is that it allows these students to be served under IDEA.

Another advantage to labeling students with disabilities is that it ensures that these students receive the proper types of assessment and the proper types of educational placements. A third advantage to labeling students with disabilities is that it ensures that these students are provided protection under the law and that they continue to receive protection under the other laws that afford rights to individuals with disabilities (e.g., the Americans with Disabilities Act) after they leave school. Lastly, some students and their families actually welcome disability labels because they provide an explanation for issues that these children and their families may have been struggling with for a long time.

There are several disadvantages to labeling students with disabilities. First, labeling a child may cause segregation of that child from his or her peers without disability labels. Placing a label on a child leads others to treat that child according to the label rather than as an individual. Labels cause stigmas that can lead to decreased self-esteem. Labels can also lead to the self-fulfilling prophecy, which occurs when low expectations lead to lower performance. When lower expectations are placed on students, those students are never challenged and so only learn to perform to lower standards. On the other hand, setting high expectations and challenging students can lead to greater achievement.

Person-centered language refers to the way in which students with disabilities are described. Placing the disability label first leads to identification of the disability first and the individual second. Instead, it is correct to refer to the individual with a disability first and the disability label second. For example, one should say "student with a learning disability" instead of "learning-disabled student." Focusing on the disability label causes others to view the student only in terms of his or her deficits or areas of weakness. If people focus on the individual first, they are more likely to see ability first and disability second.

4. ASSESSMENT OF EXCEPTIONAL STUDENTS

Under IDEA, students are entitled to nondiscriminatory evaluation. This means that they are evaluated using a variety of measures, including both formal and informal assessments. Students with disabilities should participate in the same assessments as their nondisabled peers to the maximum extent possible. This means that many students with special needs are required to participate in statewide assessments such as standardized achievement tests. Exceptional students may be easily identified by their scores on standardized tests such as intelligence quotient (IQ) tests. The bell-shaped curve refers to the distribution of scores from such tests. The majority of students fall in the middle of the distribution or the average range. Very few students fall at the extreme low or high ends of the spectrum or outside of the normal range. Students who fall at the extreme high end of the spectrum are typically students who are gifted or talented, whereas students who fall at the extreme low end of the scale are typically students with cognitive delays.

Not surprisingly, students with disabilities often perform poorly on standardized tests. There are several alternative assessment options available for students with disabilities. One of these is criterion-referenced assessment. Some commercial criterion-referenced assessments exist (e.g., Sparrow and colleagues' Vineland Adaptive Behavior Scales), whereas many are teacher-made. A criterion-referenced assessment consists of an inventory or hierarchy of repertoires across several domains, including social skills, communication skills, academic skills, leisure skills, self-help skills, and maladaptive behavior. These inventories provide a way of identifying which repertoires a student has mastered and which skills need to be taught. Identification of deficits leads to easier selection of goals. Criterion-referenced assessments are time-consuming because they may contain hundreds of items that need to be assessed with individual students, but the end result is a clear picture of the student's current skill levels.

Portfolio assessments include several different components, including samples of the student's work, progress reports and evaluations, and (sometimes) student self-evaluations. Portfolios demonstrate a student's progress over time in that samples of the student's work or tests are taken over the course of a school year or several years. For example, a writing sample taken in June may be substantially better than one taken in September. Videotapes or audiotapes may also be included in portfolio assessments.

Curriculum assessments involve teacher-made tests or assessments that come with commercially published curricula. Curricular assessments allow real measurement of student progress through items that have been taught directly over the course of the school year.

Other forms of assessment may be extremely valuable, although they are considered to be "informal" assessments. These may include reviews of records and past progress reports, observations of the student in the

classroom environment, anecdotal reports, and interviews with parents, teachers, and even the student himself or herself. Informal assessments can be valuable in obtaining information that could not be obtained from standardized or criterion-referenced tests.

5. CHARACTERISTICS OF EXCEPTIONAL STUDENTS

Exceptional students include students across 13 disability categories and students who are gifted or talented. Although some disabilities are diagnosed based solely on behavioral characteristics, no two students are alike. However, it is possible to give some general overviews of student characteristics across three broad categories: learning, social, and behavior.

5.1. Learning Characteristics

Some exceptional students excel in certain areas or learn better through certain modes of instruction. Some students, such as students with autism or mental retardation, excel in the areas of memory and fact retention but are severely deficient in the areas of language and social skills. Similarly, students with specific learning disabilities may experience difficulty in one area (e.g., verbal reasoning) but excel in another (e.g., math). Gifted and talented students and students with certain disabilities may possess one or more of what Gardner identified as multiple intelligences (e.g., bodily/kinesthetic, logical/mathematical, linguistic, spatial, musical, interpersonal, intrapersonal, naturalistic). Some exceptional students are visual learners, others learn kinesthetically, and still others learn through multiple modes of instruction. It is recommended that instruction always be presented in several ways and that students be exposed to various instructional methods, including hands-on activities, to enhance learning.

On the other hand, many exceptional students experience severe learning difficulties. Possible sources of learning problems may result from attention and memory difficulties, motivational difficulties, and/or organizational problems. Still other problems may result from a lack of generalization. Generalization can fail to occur across settings (e.g., from school to home), across individuals (e.g., from teachers to parents), and across stimuli (e.g., instructional materials presented in different formats). Teachers must plan for generalization and teach across settings, individuals, and stimuli to ensure

that responses learned under one condition are emitted under other conditions. There are several strategies available to remedy learning problems after the source of the problems has been identified.

Although some learning problems may result directly from a cognitive delay, as is the case with students with mental retardation and students with autism, other learning problems may occur due to other factors. Behavioral problems may impede learning but may also result from inappropriate instruction. If the task presented to a student is too difficult or too easy, if the pace of instruction is too fast or too slow, or if the size of the task is too large or too small, frustration or boredom may result. Frustration and boredom may lead to behavior problems. Gifted and talented students may also emit inappropriate behaviors due to boredom in their classes. Teachers must always ensure that these students have the prerequisite skills to perform the task under instruction and that they are properly motivated. It is important to figure out the learning styles of exceptional students to ensure that instruction will maximize their learning.

Because the learning styles and characteristics of exceptional students are so diverse, the IEP becomes a critical document for ensuring that students are receiving the specialized and individualized instruction they need. Certain instructional modifications that have been useful in improving learning will be outlined in the IEP to assist teachers in providing the best education possible to their exceptional students. Sometimes, however, exceptional students experience poor academic performance as a result of other factors such as social or behavioral problems that impede learning.

5.2. Social Characteristics

Exceptional students may also experience difficulties with social interactions. They may appear awkward in social situations, have difficulty in making friends, and respond differently to social situations. Often, exceptional students may have the language skills to communicate but lack knowledge of the rules of conversation such as body language, proximity, turn taking, and staying on topic. The conversational skills of students with disabilities might seem egocentric or one-sided. These students might not listen to the responses of other students or might not play the role of both speaker and listener in conversations. Exceptional students often do not learn from appropriate modeling, do not pick up subtle cues, and do not interpret the feelings of

others appropriately. Gifted or talented students may also emit awkward social behavior, perhaps because their level of maturity is above that of their same-aged peers or because they have not had experiences similar to those of their peers. Difficulties with social skills may also lead to behavior problems.

5.3. Behavioral Characteristics

Behavior problems are probably most common among exceptional students. Behavior problems may range from assaultive behaviors (e.g., hitting, screaming, throwing objects) to severe withdrawal. Behavioral issues may also include inappropriate or bizarre behaviors (e.g., stereotypy, self-talk, self-injurious behaviors). The important thing for educators and parents to realize is that all behavior has a purpose or function. Sometimes, bad behavior is the result of an inability to communicate due to a lack of language skills. Other bad behaviors may result from attempts to gain attention or to escape or avoid an aversive task. Still other behaviors (e.g., stereotypy) are emitted because the behaviors themselves are reinforcing (e.g., sensory stimulation).

Attention-seeking and escape behaviors are often mistaken for willful disobedience or are seen as a result of the disability itself. Often, fictitious explanations of behavior result from placing the locus of control within the child (or within the parents) instead of looking to environmental variables to explain behavior. From a behavioral standpoint, all behavior is controlled by its consequences. This means that a student will continue to emit only behaviors that are reinforced or that are followed by desirable consequences. If the child screams in an attempt to gain attention and an adult responds to the child's screams, the child is more likely to scream to gain attention in the future. The consequences that educators and parents deliver, such as telling the screaming student to be quiet or asking the student what he or she wants, do not serve the purposes that were intended. Attention comes in many forms, including speaking to the child and making eye contact with the child. Escape behaviors are maintained by consequences that allow the child to get out of doing the task.

Functional behavior assessment involves determining what the function or purpose of the behavior is or what consequences are maintaining the behavior. Functional behavior assessments also seek to identify the antecedents or precursors for behavior as well as environmental and motivational variables. Behavior that continues to occur is behavior that is reinforced. The function of the behavior may be communication,

attention seeking, escape, or sensory reinforcement. Under IDEA, functional behavior assessments must be conducted for all aberrant behavior. Behavior plans then become part of the student's IEP.

6. IMPLICATIONS FOR PROFESSIONALS AND FAMILIES

It is very important for educators and parents to remember that undesirable behavior always has a function. If the function of the behavior can be determined, a solution to the problem behavior is usually suggested. Blame for the behavior should not be placed on the child or on the child's disability. Although the learning, social, and behavior difficulties of exceptional children might seem overwhelming, there is a great deal of research in the field, especially in applied behavior analysis, that demonstrates successful strategies and tactics for remedying bad behavior and teaching important skills. Sometimes, the most important skills for exceptional children to learn are how to manage their own behavior, how to seek help when necessary, and how to advocate for themselves. Fortunately, under IDEA, exceptional children have access to a wealth of pedagogical practices designed to prepare them to live as independently as possible.

See Also the Following Articles

Educational Achievement and Culture ■ Gifted Students ■ Learning ■ Mathematics, Teaching of ■ Reading, Teaching of ■ Special Education ■ Teaching Effectiveness ■ Writing, Teaching of

Further Reading

Alberto, P. A., & Troutman, A. C. (2003). *Applied behavior analysis for teachers* (6th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Bos, C. S., & Vaughn, S. (2002). *Strategies for teaching students with learning and behavior problems* (5th ed.). Boston: Allyn & Bacon.

Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis.* New York: Macmillan.

Friend, M., & Bursuck, W. D. (2002). *Including students with special needs* (3rd ed.). Boston: Allyn & Bacon.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century.* New York: Basic Books.

Heward, W. L. (2003). *Exceptional children: An introduction to special education* (7th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Hunt, N., & Marshall, K. (2002). *Exceptional children and youth*. Boston: Houghton Mifflin.

Lewis, R. B., & Doorlag, D. H. (2003). *Teaching special students in the general education classroom* (6th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Lipsky, D., & Gartner, A. (Eds.). (1989). *Beyond separate education: Quality education for all*. Baltimore, MD: Brookes.

Mastropieri, M. A., & Scruggs, T. E. (2004). *The inclusive classroom: Strategies for effective inclusion* (2nd ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (2003). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.

Turnbull, R., Turnbull, A., Shank, M., & Leal, S. J. (2004). *Exceptional lives: Special education in today's schools* (4th ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.

U.S. Department of Education, and Office of Special Programs, Data Analysis System. (2002). *To assure the free appropriate public education of all children with disabilities: Twenty-fourth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: U.S. Department of Education.

Vaughn, S., Bos, C. S., & Schumm, J. S. (2003). *Teaching exceptional, diverse, and at-risk students in the general education classroom* (3rd ed.). Boston: Allyn & Bacon.

Executive Development
and Coaching

Rob Silzer

HR Assessment and Development Inc., New York, New York, USA

GLOSSARY

coaching Helping others to improve their knowledge, skills, abilities, and/or performance.

development Learning or strengthening knowledge, skills, abilities, or performance.

executive A general manager, head of a functional area, or corporate officer in an organization.

general manager A manager who has multiple organizational function responsibilities for running a business (e.g., production, sales and marketing, product development, human resources, finance).

global executive An executive who can work effectively and comfortably across cultural borders and who typically has lived and worked in at least two different countries.

organizational culture The shared values, beliefs, practices, rituals, and behavioral patterns in an organization.

strategic thinking Understanding and setting distinctive competitive goals and actions for an organization.

Executive development and coaching includes a variety of assessment tools and interventions that have been designed to help executives improve their knowledge, skills, abilities, and performance and to build the talent needed to meet strategic objectives and new business challenges in an organization.

1. WHAT IS THE HISTORICAL CONTEXT?

A greater focus on executive development and coaching emerged during the 1990s. During prior decades, executives were not given as much attention; instead, the emphasis was on selecting individuals who had management potential and on using basic management training to teach functional skills.

During the 1980s, the world of business changed significantly. There was a significant globalization of business markets and a destabilization of many companies and industries due to increased global competition, to a wave of business mergers and consolidations, and to an expanding acceptance of management talent moving across companies. The speed of business and decisions

greatly increased and forced people to work hard to keep up with the rapid changes. This led to a need for fungible managers and executives who can learn quickly and be effective in a range of business situations. Managers who can do this are highly valued, and they often leverage this demand for talent to aggressively negotiate for career opportunities that move them into executive roles so that they can help run a company.

During the 1990s, executives emerged as the key to effective corporations. They were seen as critical decision makers and were given a lot of personal visibility and credit in the press. In addition, many managers viewed being an executive and running their own businesses as viable career goals. However, it became evident that many managers and executives did not have the skills and abilities to effectively handle an executive role or to adequately adapt to major and rapid changes occurring in the business world. Because of the increased competition for the best talent, companies aggressively pursued initiatives to better train and develop their current internal managers and executives.

Initially, these executive development initiatives focused on sending current executives and managers with executive potential to graduate business schools for executive development classes and executive MBA programs. These efforts focused on providing individuals with additional business knowledge, particularly in corporate finance. Because of the cost and the limited number of participant openings in these external programs, companies began to conduct internal executive development classes and tailoring them to the specific needs and business situation of each company. Later, these programs expanded to also include some emphasis on the development of executive performance skills such as strategic thinking, media and press interviewing, and decision-making skills.

As corporations went through broad reorganizations and rapid changes, some executives had difficulty adapting to the changes. This led to the increasing turnover of executives and chief executive officers (CEOs) in companies, which in turn placed new performance demands on the remaining executives in a company. The existing executives were expected to retune their skills to fit the new expectations and strategic directions of the recently hired CEO. Human resource officers began to hire executive development experts, including retired business executives and organizational psychologists, to work in groups and individually with these executives in order to "fix their skills." After some success in getting executives retrained and aligned with the new CEO, these development experts were then assigned to help well-performing executives and high-potential managers who wanted to continue to build their executive skills. Slowly, the field of executive coaching emerged and evolved from fixing performance problems to further building the skills of strong performers.

This article reviews some of the current techniques used in executive development and coaching. When a company considers starting an executive development initiative, it typically must consider a few fundamental issues. How a company or development professional thinks about these issues has a direct bearing on the nature, approach, and success of the development effort.

2. WHO IS AN EFFECTIVE EXECUTIVE?

There are many opinions on which knowledge, skills, and abilities (KSAs) an executive needs to perform effectively. Although some people use the term loosely, this article will reserve the term "executive" for general managers, functional area heads, and corporate officers in an organization. Many companies, human resource professionals, and industrial/organizational psychologists have identified their own lists of the necessary executive competencies. In 2002, Silzer and a range of respected experts reviewed executive skills and abilities from a variety of angles, while others, such as Fulkerson, McCall, and Hollenbeck, focused on the requirements to be an effective global executive. Silzer, in his chapter on selecting executives, outlined the challenge of selecting individuals who would be successful over the long term and across a variety of business challenges. He discussed the complexity of the selection process in terms of "executive fit" on four different levels:

- Person–position fit (matching the executive with the specific job requirements)
- Person–group fit (matching the executive with the existing executive team and direct reports)
- Person–organization fit (matching the executive with the organizational culture)
- Person–cultural fit (matching the executive with the country culture)

Some people argue that the executive job is so complicated, individually tailored, and ever changing that it is impossible or useless to identify one set of KSAs that apply across all executives and all situations. Unfortunately, this view leads down an empty path and provides little help when trying to structure

and implement executive selection and development approaches. Others would argue that while the executive role is very complex, there are some fundamental performance areas that do apply to a wide range of executives and business situations. A list of key executive KSAs that are relevant to most executive positions probably would include the following:

- Sound judgment and decision making
- Strategic thinking and strategy management
- Interpersonal influence
- Communication effectiveness
- Leadership confidence and impact
- Results orientation
- Business and financial acumen
- Customer focus
- Integrity
- Organizational insight and management

Executives are also typically expected to have some knowledge of general business practices, functional areas, organizations, financial management, and human resource management. Lately, knowledge of information technology, marketing, and customer service has become more important to executives, whereas knowledge of particular industries or products has become less important because executives now frequently switch companies and industries during their careers.

Of course, executives also need to handle their own, ongoing job responsibilities. These vary widely depending on the specific executive role, but there are some performance areas that apply to a broad range of executive positions, including the following:

- Accomplishing annual goals
- Managing the organization and financial budgets
- Building, motivating, and developing leadership talent
- Pursuing innovative ideas and distinctive competitive strategies
- Defining and supporting the organizational culture

3. WHAT IS EXECUTIVE DEVELOPMENT?

In the past, there was a strong "selection bias" in human resource management. That is, there was a widespread belief that managers and executives had limited ability to grow or change, so executives, managers, and human resource professionals relied almost exclusively on selection to find the right person for a management or executive job. If someone was performing poorly, that person was fired or moved aside so that another person could be selected to fill the position. Most psychological assessment tools and techniques in organizations were used to select the best person for the jobs.

However, during the 1980s, the field of training and development started to emerge as useful and relevant to management and executive effectiveness. For example, assessment centers that previously had been used exclusively for selection were being used to develop current and future incumbents to be more effective in these roles. Particularly in the fast-changing business environment of the 1980s and 1990s, companies had to figure out how to quickly redeploy and redevelop current employees for new roles and jobs because it often was infeasible to fire everyone and hire an entirely new workforce and executive team (although some companies did exactly that). As a result, there was a wave of interest in management, leadership, and executive development techniques.

The term "executive development" means learning or strengthening the KSAs needed to be effective in an executive position. There was significant interest by organizations in identifying and further developing those individuals who were seen as having significant potential to be effective in higher level executive positions. This led to the development of a range of new initiatives, including succession management systems, executive classes and programs based in graduate business schools, and the widespread use of management and executive competency models to specify the KSAs that a manager or an executive needs to be effective.

4. WHAT ARE TYPICAL EXECUTIVE DEVELOPMENT APPROACHES?

Companies and development professionals have devoted a good deal of attention and creativity to designing management and executive development programs. In 1999, Conger and Benjamin outlined some development approaches used by different companies. In thinking about executive development, it is helpful to separate the approaches that focus on the assessment of executive KSAs and performance from those techniques that are used as development interventions.

4.1. Assessment Approaches

Some executive development programs are designed so that all executives, regardless of their individual skills

and performance, go through the same program either as individuals or as groups. Another approach that gained support and acceptance during the 1980s is to first identify the development needs of individual executives and to provide a development experience that is appropriate for those specific needs. This leads to an expanding role for assessment in the process.

Assessment of an executive's development needs has typically used one of three general approaches:

- Performance reviews
- Multisource feedback
- Individual (psychological) assessment

4.1.1. Performance Reviews

The most common assessment approach is to use the annual performance review to identify an individual's performance/KSA strengths and development needs. Nearly every organization has designed an annual executive evaluation process that reviews an executive's behavior and performance in several areas that may include looking at results against annual goals, general executive performance, executive skills and abilities, and how well the executive demonstrated company values. Some review programs include specific behavioral standards of expected performance as review guidelines, whereas others are more open-ended and rely on the broad judgment of the reviewer.

4.1.2. Multisource Feedback

Another assessment technique that has gained widespread acceptance is a multisource feedback approach. This relies on collecting structured feedback on an executive's performance and KSAs from a range of people who have had an opportunity to interact with the executive and to observe the executive's behavior. This group might include immediate or higher level bosses, peers, direct and second-level reports, internal and external customers, members of the board of directors, outside community and financial contacts, vendors, and suppliers. The feedback is typically collected either through a structured written questionnaire (often referred to as a 360-degree feedback instrument), telephone interviews, or one-on-one meetings. A written questionnaire provides a more standardized approach that can cover quite a broad range of performance areas and KSAs. The interview and meeting approaches allow for a deeper discussion of key issues and the identification of numerous examples of behavior that illustrate the feedback.

4.1.3. Individual Assessment

A third approach that has been gaining significant support is an individual (psychological) assessment of the executive. This approach was initially used by companies to help select executives but is now also being used extensively for development. In 1998, Jeanneret and Silzer provided a good resource on how psychological assessment tools and approaches can be used effectively in organizations. When used for development, this approach is typically designed to assess current performance, KSAs, development needs, career direction, and potential for other roles. The approach usually includes a range of tools and techniques such as the following:

- Cognitive ability tests
- Personality and career interest inventories
- Structured behavioral interviews
- Behavioral exercises and simulations (e.g., role-play, "in" basket, and presentation exercises)
- Multisource feedback

Assessment techniques are usually used with individuals, although it is becoming more common for all executives to go through an assessment process at the same time. This group approach not only allows the company to understand the strengths and weaknesses of each individual executive but also provides a profile of the entire executive group. The CEO, and often the board of directors, can then see the shared strengths and shortcomings of the group. This approach can be very useful when the organization is going through a significant change such as a merger, downsizing, reorganization, or change in strategic direction. It can lead to a valuable reshuffling of executive talent across the organization to better match individuals to executive roles. In addition, the process can identify common performance weaknesses that can then lead to a group development initiative or even a modification in the company's strategic goals. For example, in some cases, organizations have used an executive group profile to determine the strategic direction of the companies by looking for opportunities that leverage the strengths of the existing executives and that minimize any reliance on areas that are group weaknesses.

4.2. Intervention Approaches

Once the development needs or goals have been identified, then appropriate development intervention approaches can be identified that specifically address those needs. There are myriad intervention tools and

techniques that have been used to develop executives. They might be categorized into four general types:

- Learning through the experience of others as an individual.
- Learning through the experience of others with a group.
- Learning by personal experience as an individual.
- Learning by personal experience with a group.

4.2.1. Learning through the Experience of Others as an Individual

Initially, this was probably the most common development approach, with individuals learning from their immediate bosses or the previous incumbents in their roles. When business started to go through widespread change, many bosses and previous incumbents no longer were the best sources for the executives on how to handle the new challenges. So, executives started turning to others with more experience for help. Some approaches to executive development in this category include the following:

- Working with an internal or external mentor
- Locating a tutor or an expert on a specific topic
- Establishing an internal or external network of resources and experts
- Taking a temporary assignment with a customer or vendor
- Initiating a personal reading program
- Conducting a study or writing a white paper on a critical business issue
- Visiting competitors, suppliers, and customers
- Attending conferences and seminars
- Teaching a business class

4.2.2. Learning through the Experience of Others with a Group

This has been a frequently used approach, often relying on attending business classes or executive programs at graduate business schools. The high cost of these programs limited the number of executives who could attend and eventually led to a decrease in attendance but also to an increase in internal executive programs designed to meet the same goals. A group approach to development typically means that all executives go through the same program without regard for their individual development needs. However, in some cases, executives with similar needs have joined together to address those specific common needs as a group. In 2002, Marcos described some innovative group approaches. Some examples of group development include the following:

- Attending a business school course or an executive program
- Participating in an internal leadership or executive training program
- Starting a reading discussion group
- Participating in a peer group shared learning program
- Initiating an internal speaker presentation series

4.2.3. Learning by Personal Experience as an Individual

As the executive development field moves to assessing and identifying individual development needs, there is a growing focus on activities in this category. Many development action plans are being written that focus on the specific development needs of the executive. These plans initially and primarily consisted of taking courses, but the field soon adopted the now widely accepted idea that the best development occurs on the job in a "learn-by-doing" approach. This shifted quite a bit of attention to experiential learning at the executive level. One very visible technique is the "action learning" approach described by Dotlich and Noel in 1998. Other examples include the following:

- Starting an on-the-job skills building program
- Initiating an action learning assignment and project
- Rotating through different core businesses or functional areas in the company
- Becoming an executive assistant to an officer
- Leading a strategy review or a negotiating process
- Taking an overseas or global assignment
- Managing a start-up, turnaround, or high-growth business
- Managing a large, highly complex, or matrixed business
- Taking responsibility for a liaison role such as government lobbying or community affairs

4.2.4. Learning by Personal Experience with a Group

Some of the preceding approaches evolved into group experiences. In some companies, "action learning" means group projects. A selected team of executives or high-potential managers is assigned to work on a

special project that addresses a major organizational issue or problem. The team learns a great deal about the issue and the organization, while the company gains some sound recommendations on how to address the issue. Examples of group learning approaches include the following:

- Joining an action learning project team
- Volunteering for a corporate task force
- Being part of a new business start-up team
- Pursuing an internal or external business partnership
- Joining an industry-wide committee
- Serving on a business review team

5. WHAT IS EXECUTIVE COACHING?

During the past 10 years, executive coaching has become a widespread executive development tool. It can be defined as a one-on-one helping relationship between an executive and a coach that helps to improve the executive's performance and combines learning through the experience of others with learning from personal experience. Often, this process focuses on the executive's KSAs and behavior. It is different from mentoring in that coaching seeks to change the executive's on-the-job behavior and performance and not just to provide advice. Hollenbeck, in his enlightening 2002 chapter on executive coaching, distinguished it from psychotherapy by noting that the goal of coaching is "better job performance," whereas the goal of psychotherapy is "a more fully functioning person."

In some cases, every executive in a company is assigned a personal external coach. A number of companies later backed off of this sweeping approach due to the high cost. Executive coaching turned into something of a fad and attracted the interest of a wide range of coaching service providers. Because there are no credential standards, anyone can assume the title of executive coach. This includes human resource consultants, trainers, psychologists, retired business executives, health care providers, and MBA graduates. As a result, the field of executive coaching and executive development has been swamped with people who do not have the necessary coaching skills or business knowledge. Corporations are often fairly indiscriminate in selecting executive coaches. Many coaches are probably ineffective, but corporations do not spend much time in evaluating the effectiveness of coaching interventions. Under significant cost reduction pressures, companies are now starting to better manage these assignments. This has resulted in a few emerging trends in executive coaching, including the following:

- A more discriminating use of external coaches
- Written standards for the selection and evaluation of coaches
- The training of current internal executives as peer coaches

Good coaches need to have knowledge of business and executive performance issues as well as skills in communicating, interpersonal influencing, and changing behavior. Many coaches fall far short of this. The coaching process typically lasts from 6 to 18 months, although many companies have started to put a 1-year time limit on these interventions. The process generally consists of five phases:

- Initial agreement and contract
- Assessment
- Feedback and planning
- Coaching intervention
- Follow-up

Moving through these phases depends on success at each step. Coaches need to be able to demonstrate progress in changing an executive's behavior. Too many interventions are just pleasant bimonthly conversations with little focus or impact. Coaches must also be trustworthy and be able to carefully manage the confidential information shared by the executive and the executive's immediate bosses. Many coaching interventions fail because of a poor coaching contract at the beginning of the process, a lack of confidentiality, or a lack of change in the executive's behavior.

Recent interest in executive coaching has focused on a few key issues:

- How can the process be made more cost-effective?
- How can the effectiveness of coaching interventions be rigorously evaluated?
- How can individual coaches be objectively selected, assigned, and evaluated?

6. WHICH PERFORMANCE AREAS ARE MORE AMENABLE TO DEVELOPMENT?

Over the years, coaching experts have gained a good deal of insight into how to effectively change the

behavior and performance of executives. There is some agreement that interpersonal and communication skills are the most common areas that coaches focus on for development. They also are the areas where good coaches can have significant impact. Other areas that seem fairly amenable to development include personal organization, planning, administrative skills, business knowledge, presentation skills, financial knowledge, and project management.

On the other hand, there are a number of areas that are difficult to develop either due to the complexity of the area or because it requires specific preliminary abilities. These areas include complex decision making, innovative thinking, strategic planning, personal and career motivation, leadership versatility and influence, and personal integrity.

The field still needs to better define which development approaches are best suited for which performance areas. It is now known that interpersonal role-playing and behavioral exercises do seem to be useful approaches in changing an executive's interpersonal skills. However, there is little research matching development interventions with other performance areas. In general, an executive needs to go through four stages for changing behavior: understanding the behavioral issue, learning new effective behaviors, practicing the new behaviors, and getting feedback from others on those behavioral attempts.

7. WHICH EXECUTIVES ARE MORE LIKELY TO DEVELOP?

There also has been growing agreement that some executives are more likely to develop than others. Although there is not total agreement on what separates these two groups, there are a few markers that seem to distinguish the developable executives. These executives generally have strong cognitive skills in order to adequately understand the development that is needed. They also need to be mentally flexible and behaviorally adaptive to try new behaviors, and they should have a clear motivation and willingness to change. That is, they should see the need to change their behavior. In addition, it helps if they have a learning orientation, that is, an intellectual curiosity and interest in learning new things. Executives who do not develop generally are inflexible, unwilling to change, or unable to learn new behaviors.

8. WHAT ARE THE HALLMARKS OF DEVELOPMENT SUCCESS AND FAILURE?

Companies typically spend significant financial resources on executive development each year. There is a large group of internal and external development specialists and consultants who have been very successful in providing executive development services. Although many companies have recently worked to reduce development costs, many expensive development approaches, such as business school programs and executive coaching, are still widely used.

Executive development efforts often fail to have any impact on an executive, although companies rarely hold these programs accountable for producing measurable results. It may be some time before companies implement tough outcome standards. There are some known reasons why many of these efforts fail, including the following:

- Lack of clear objectives
- Inappropriate or irrelevant development activities
- Lack of follow through on a development plan
- Failure of the executive to learn the right behavior or lessons
- Inability of the executive to learn or change
- Lack of leadership support in the organization for the change
- Lack of business application for the change

At the same time, there are some hallmarks of successful development efforts, including the following:

- An executive who is able and motivated to learn
- Use of the right approaches and techniques for the development need
- Clear learning and behavioral objectives
- Hands-on learning experience
- Focus on specific behaviors and outcomes
- Clear connection to business applications
- Regular feedback and guidance from others
- Support from senior leaders

The emphasis on executive development is likely to continue and even increase in the future as the globalization of business increases. The demand for effective fungible executives will remain high, and corporations will see the benefit of having well-integrated executive selection and development efforts to meet new business challenges and to achieve difficult strategic objectives.

See Also the Following Articles

Organizational Culture and Climate ∎ Person–Environment Fit ∎ Work Teams

Further Reading

Conger, J. A., & Benjamin, B. (1999). *Building leaders: How successful companies develop the next generation.* San Francisco: Jossey–Bass.

Dotlich, D. L., & Noel, J. L. (1998). *Action learning: How the world's top companies are re-creating their leaders and themselves.* San Francisco: Jossey–Bass.

Fulkerson, J. (2002). Growing global executives. In R. Silzer (Ed.), *The 21st century executive: Innovative practices for growing leaders at the top* (pp. 300–334). San Francisco: Jossey–Bass.

Hollenbeck, G. (2002). Coaching executives. In R. Silzer (Ed.), *The 21st century executive: Innovative practices for growing leaders at the top* (pp. 137–167). San Francisco: Jossey–Bass.

Jeanneret, R., & Silzer, R. (1998). *Individual psychological assessment: Predicting behavior in organizational settings.* San Francisco: Jossey–Bass.

Marcos, V. (2002). Developing executives. In R. Silzer (Ed.), *The 21st century executive: Innovative practices for growing leaders at the top* (pp. 168–199). San Francisco: Jossey–Bass.

McCall, M., & Hollenbeck, G. (2002). *Developing executives: The lessons of international experience.* Boston: Harvard Business School Press.

Silzer, R. (2002). Selecting leaders at the top: Exploring the complexity of executive fit. In R. Silzer (Ed.), *The 21st century executive: Innovative practices for growing leaders at the top* (pp. 77–113). San Francisco: Jossey–Bass.

Silzer, R. (Ed.). (2002). *The 21st century executive: Innovative practices for growing leaders at the top.* San Francisco: Jossey–Bass.

Extreme Environments and Mental Function

Robert B. Bechtel

University of Arizona, Tucson, Arizona, USA

GLOSSARY

Altiplano The highlands of Peru.

chacheeko Local name for a newcomer in Alaska.

cold soaked Frozen.

dehydration Loss of water in the body.

desertification Land turning into a desert.

diurnal rhythms The daily cycle of the sun.

head bolt An electrical outlet for cars to plug into at a parking lot.

heat exhaustion Collapse due to overheating and loss of water.

heatstroke More serious condition due to body's failure to deal with heat.

hypothermia Loss of body heat.

hypoxia Lack of oxygen.

Monge's disease Altitude sickness.

photophobia fear of light.

puna Local word in Bolivia for altitude sickness.

soroche Local word in Peru for altitude sickness.

tacya Ironwood digging stick used in the Peruvian Altiplano.

the 200 Club Initiation rite at the U.S. polar station; one joins by running naked to the pole and back when the temperature is −100° F.

Extreme environments are of two kinds: the permanent ones where people live year-round (e.g., high mountains, deserts, cold regions, jungles) and the temporary ones such as space flight, simulation laboratories, and the temporary incursions into the permanent extremes (e.g., mountain climbers, missionaries, explorers). Problems of function are most frequent in the temporary incursions into the permanent areas and places such as Antarctica, space flight, and the simulation laboratories. Extremes are being made more endurable by introducing communications technology and more creature comforts.

1. DEFINITIONS

The definition of extreme environments is very geocentered. People living in the heavily populated and industrialized latitudes around the tropics of Capricorn and Cancer assume that equatorial and arctic climes are extreme, whereas those living in these "extreme" latitudes think of them as normal. Similarly, those near sea level think of altitudes of a mile or more as extreme, whereas those living at these altitudes think of them as healthier. Thus, one person's extreme can be another person's normal surrounding. "Extreme" as a label acquires some credence when it involves environmental conditions that bring about physiological responses that

interfere with the normal functioning required to get the business of living done. In high altitudes, this includes hypoxia (i.e., lack of oxygen); in colder climates, it includes frostbite and hypothermia; and in hot climates, it includes heat stroke, heat exhaustion, and dehydration. In jungles, what it includes is less clear, but extreme humidity, exotic diseases, and a myriad of insects make it at least unusual compared with more temperate climes.

When it comes to environments such as space travel, balloon exploration, and experimental laboratories that simulate space travel, no one has any trouble calling these extreme. Similarly, when mountain climbers, explorers, and skiers climb to great heights, these are also seen as extreme.

A distinction must be made, however, between those latitudes and altitudes where people live year-round and have a chance to adapt through generations and those extremes where visitors enter for a brief time and then leave. The latitudes and altitudes where people live year-round under harsh conditions are a permanent situation, whereas those where there are only short incursions are best labeled "temporary."

2. THE ENVIRONMENTS

2.1. The Temporary Environments

2.1.1. Antarctica

Although Antarctica is now inhabited year-round by people from a variety of countries, all of these people are there on assignment and will rotate home at some point. No one has been born in Antarctica, although some have died there (largely by accident). Antarctica has been characterized as the highest, driest, and coldest desert on earth. No one thinks of it as a desert due to the miles of ice that have accumulated there in glaciers, but its average precipitation makes it comparable to deserts at more northern latitudes. Thus, Antarctica combines all of the climatic elements that make environments extreme—cold, altitude, unusual diurnal rhythms, and dryness—in addition to another element: isolation.

At an average altitude of 9000 feet (2743 meters), many people experience hypoxia (i.e., lack of oxygen). Hypoxia can cause sleep disturbances, and this can lead to loss of memory, hallucinations, and other mental aberrations. Therefore, sleep deprivation alone could account for failure to perform duties. The average person who goes there, however, experiences less

dramatic symptoms such as failure to attend properly at random times and the resultant mistakes.

The diurnal effects on sleep are also notorious. When Burgess Ledbetter and the author were doing their work in Alaska, they encountered a man who had not slept in 10 days. Thus, the effects of long daylight, as well as the long, dark, and cold nights, can be felt.

The cold of Antarctica has a dramatic effect. Jerri Nielson, the doctor who was rescued from Antarctica due to her cancer, described how the impact of the cold made her forget the orientation lecture. Rivolier, in his field studies, found that the impact of even minimal exposure to the cold was measurable. One reaction that affects performance is the extension of reaction time. This is especially troublesome in outside activities such as driving vehicles and repairing. Perhaps the most common malaise reported in Antarctica is the semicomatose state referred to by many names such as "cabin fever," being "stir crazy," having the "big eye," and just "staring into space." Not everyone experiences or reports this, but nearly everyone reports periods of wandering attention. The lack of attention can be serious when a schedule must be followed, and there are many reports of personnel who have let schedules slip.

Often overlooked in the concentration on negative experiences are the more positive outcomes reported by many inhabitants of the Antarctic stations. For some, the winter in Antarctica is the "experience of a lifetime." Nielsen pointed out that Antarctica can be habit forming.

Finally, there is the form of madness called "the 200 Club." This rite of initiation occurs at the pole station when the temperature drops to –100° F (–73.3° C). The initiates are heated up in a room at +100° F (+37.7° C) and then run naked (with shoes on) to the pole and back. The spectacle is commemorated with photos of icicles dripping from appendages.

2.1.2. Space Travel

Most of the experience in space travel comes from the Russian space station *Mir,* the space shuttle, and the exploratory *Apollo* missions to the moon. The greatest danger in space travel, radiation from the sun, has largely been avoided by rotating the vessels, traveling at times of the quiet sun, and experiencing just plain good luck. The second most serious problem has not actually been dealt with in a remedial fashion. Weightlessness causes the body to dissolve calcium out of the bones. This can cause a serious disability such as osteoporosis. Russian cosmonauts who spent a

long time on *Mir* had to be carried out of the return capsule and took weeks to recover sufficiently to walk normally again. A program of 6 to 8 hours of strong exercise aboard *Mir* did not help alleviate the problem. There seem to be large individual differences in the loss of calcium. Some astronauts were able to walk after landing, whereas others were not. There remains a question as to whether the bone density ever returns to normal. Although bones have not been broken due to this problem so far, it would be a serious problem for longer voyages such as one to Mars.

The trip to Mars constitutes the most extreme of any environment to date. This is due to the communications lag between Earth and Mars of up to 40 minutes and the length of time spent traveling to, staying on, and returning from Mars that could take up to 3 years. Because it is not known whether there is any water on Mars, urine will have to be recycled through a purifying filter (Fig. 1). The "home" on Mars will consist of a reused rocket body (Fig. 2). A human factors problem is the need for windows that considerably weaken the structural integrity of the rocket body.

As in other isolated confined environments, space travel wreaks havoc on sleep patterns. Astronauts and cosmonauts can lose 2 hours of sleep every 24 hours and end up with serious sleep deprivation. Although this has not resulted in hallucinations or psychotic behavior, the lack of attention to detail and time can result in serious mistakes.

2.1.3. Simulation Laboratories

In September 1991, eight biospherians were confined in a Mars colony simulation and remained there for 2 years (Fig. 3). Before the 2 years were up, oxygen had to be pumped into the biosphere structure and higher calorie foods were introduced. The atmosphere was reduced to the equivalent of 13,000 feet (3962 meters), and the number of calories in the diet had fallen to 1800 per day. In addition, a conflict had developed among the crew members. This experience was similar to those of Russian attempts at simulation. Nevertheless, the crew finished the mission. Much earlier, the Tektite Lab, an underwater laboratory, had uncomfortable conditions of crowding, humidity, and disturbed sleep. But this crew also finished the mission and did even more work than is done under normal conditions.

The experiences with the simulation laboratories are mixed. Some result in conflict, whereas others do not. The simulations are often considered to be good training.
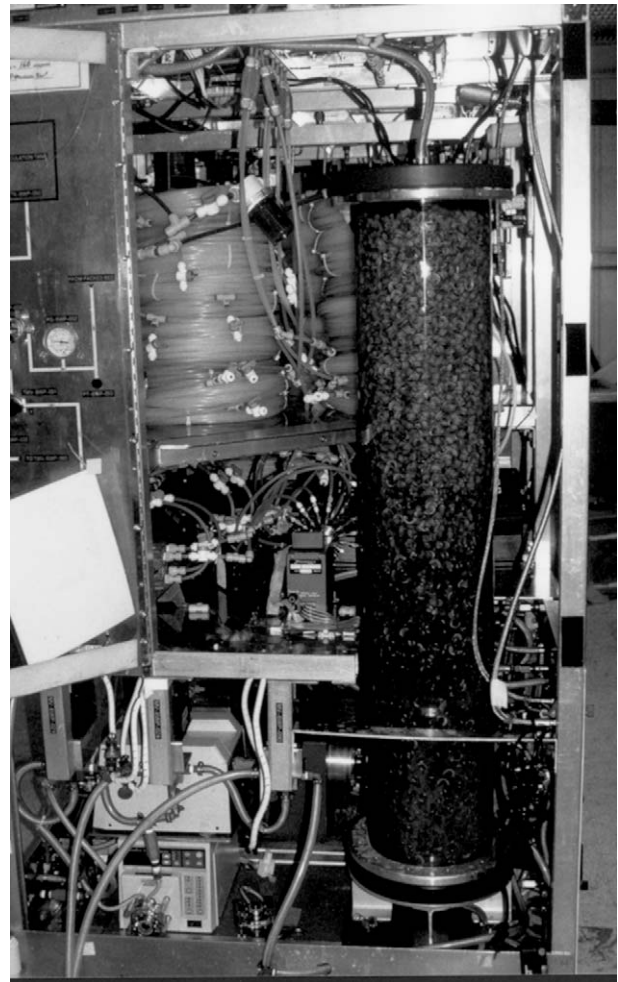

FIGURE 1 Urine recycling apparatus at the National Aeronautical and Space Administration (NASA).

2.2. Permanent Extreme Environments

2.2.1. High Mountains

There are really only three areas on Earth where significant numbers of people live permanently at high altitudes: the Tibetan Plateau, the Andes, and the highlands of Ethiopia. These are also areas where mountain climbers and miners sometimes work. The continent of Antarctica does not qualify because all persons there are on temporary assignment. Only visitors to these places have the kind of symptoms that seriously interfere with mental function. Lack of oxygen is the primary environmental element that leads to interference with normal functioning (hypoxia) at high altitudes. There are two types of altitude sickness: acute mountain sickness and

FIGURE 2 Rocket body simulating crew quarters for Mars.
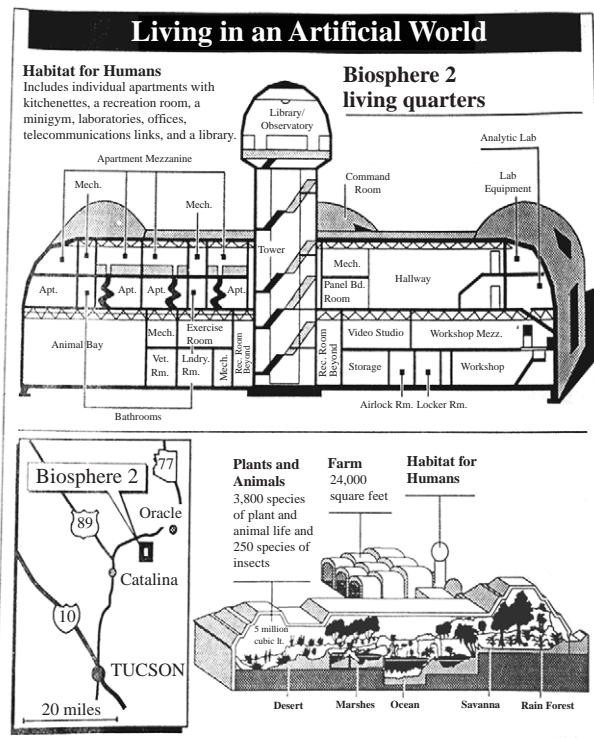
FIGURE 3 Diagram of Biosphere 2. Apt., apartment; Bd, board; Mech, mechanical; Vet, veterinary.

chronic mountain sickness. Acute altitude sickness, which is what visitors experience, is also known as Monge's disease; it is also known locally as *soroche* in Peru and as *puna* in Bolivia. The person with hypoxia experiences restless nights, headache, vomiting (sometimes), tightness in the chest, irritability, and a fear of light (photophobia). Acute altitude sickness can also produce hallucinations and erratic behavior. There are many tales of this kind of behavior. One woman, a native of Peru, described her companion's constant efforts to turn off the car lights while driving on a mountain road. It was the most frightening ride of her life. Another traveler to the high mountains told of his wife's description of a blue horse. These symptoms usually disappear by the fifth day.

Chronic mountain sickness involves swelling of the brain and lungs. This can lead to death, and people who suffer from the symptoms must be quickly removed and taken to lower altitudes. The military has established a rule that oxygen masks should be used above 10,000 feet (3048 meters). Mount Everest is more than 29,000 feet high (8839 meters), yet in 1978 Reinhold Messner climbed to the top without supplemental oxygen. He went on to do it alone in 1980 and then, again without supplementary oxygen, climbed all 14 mountains in the world that are more

than 8000 meters high. This represented another monument to human adaptability.

2.2.2. Deserts

Deserts are characterized as being hot, dry, and devoid of the same kind of plant and animal life found in more temperate areas. Authorities disagree on definitions of deserts. Some prefer an extreme of 2 inches or less of rainfall per annum. Others' definitions go up to 15 inches of rain per annum. Some compromise and call these "arid lands." The definition of an arid land can be any place that requires irrigation. Areas on Earth that are seen as deserts include places such as the Sahara in Africa (the largest), the Namib in southern Africa, the Atacama in South America, and Death Valley in the southwestern United States. Areas such as the Sonoran Desert in the southwestern United States and northern Mexico, the Great Sandy in Australia, the Gobi in Mongolia, and nearly all of Saudi Arabia (Fig. 4) are also some of the more obvious ones. It should also be noted again that great parts of the Arctic and Antarctica are also considered desert if desert is defined by low precipitation. The main symptoms suffered in deserts seem to be chiefly from exposure to the sun with the result of heatstroke or heat exhaustion. Both involve many of the same symptoms: faintness, confusion, fainting, and vomiting. But the treatment is very different. Heatstroke, the most serious, requires immediate cooling, whereas heat exhaustion requires immediate water intake (salt is no longer recommended). Heatstroke is best recognized by a lack of perspiration. Thus, dragging the person into the shade and cooling with a water hose or any other method are critical.

It is estimated that the longest the average person could last if lost in the desert is 2 days. Death occurs when the body temperature rises above 113° F (45° C). Having said this, it should also be remembered that in 1905, Pablo Valencia got lost in the desert near Yuma, Arizona, and survived after 6½ days. He traveled only at night and rested during the days. Although many may regard the Grand Canyon as not in the desert category, in 1996 there were 482 search and rescue operations, the result of which 18 people died. It is estimated that more than 700 people have died in the Grand Canyon, with some of these dying from freezing during the winter. Most of the distress came from dehydration. Thus, the canyon seems to fit the temporary extreme category, even though people live at the rim and at the bottom year-round.

FIGURE 4 The Rub Al Khali (The Hidden Quarter) desert in Saudi Arabia.

Recently, the Sonoran Desert claimed a large number of deaths: more than 30 in 2002 and more than 50 in 2003 due to the illegal migration from Mexico into the United States. All of these people died from dehydration.

2.2.3. Cold Regions

The far northern communities in the U.S. state of Alaska as well as those in Canada, Greenland, and the former Soviet Union are easily seen as cold regions, and for many reasons parts of Norway and Finland, all of Iceland, and even the U.S. New England states and Montana winters can seem extreme. Antarctica qualifies as well, but this continent is in a class by itself, as mentioned earlier.

Frostbite is the most immediate danger in northern climes. The military coined the phrase "30–30–30," meaning that flesh freezes in 30 seconds in a 30-mile per hour wind at –30° F (–34.4° C). Despite this, one

can see the black scabs hanging from the cheeks of many Inuit from riding their snowmobiles on the streets of many communities in Canada and Alaska. The average human body stops functioning below 77° F (25° C).

Hypothermia begins with shivering and is then replaced by muscular rigidity, confusion, and loss of consciousness. The confusion has caused many deaths, and to this day it is still sometimes reported in Alaskan newspapers that people, usually *chacheekos* (newcomers), have frozen to death while trying to walk back to gas stations that were within sight after their cars had stalled.

Having set the preceding limits by clinical data, it should be noted that in October 1992 at Point Hope, Alaska, Morris Sage and Sayers Tuzroyluck fell into a lake while fishing. Sage chose to stay in the water because it was +25° F (−3.9° C), whereas the air was −10° F (−23.3° C). Tuzroyluck chose to walk back to the town 7 miles away. Medical authorities say that the average person can last no more than approximately 30 minutes in such conditions. Both men spent 3½ hours in the cold, and Sage's body temperature was measured at 75° F (23.9° C) when he was found. Both men survived without ill effects. This is perhaps testimony to the adaptation of the Inuit people to their climate or to the tremendous variability of humans in response to extreme conditions.

Automobiles can be a hazard in cold regions. They require a complete overhaul, with all fluids (e.g., brake fluid, crankcase oil) being replaced with a thinner variety, the crankcase being wired with a heated dipstick, and the entire apparatus being wired to connect with an electric "head bolt" while parked. If not, the car will freeze—get "cold soaked" in local parlance—in approximately 45 minutes. Many choose to let the engine idle when head bolts are not available.

Diurnal rhythms are a great disrupter of sleep. Many residents of Alaska use blackout curtains and shades during the long summer daylight period to help them sleep.

2.2.4. Jungles

Perhaps too much has been claimed for jungles as extreme climates. They do not have the extreme heat of the desert, nor do they have the heights of the mountains. Yet the humidity combined with a fairly high temperature makes activity uncomfortable and can produce heat exhaustion and heat stroke under certain conditions. Diseases such as malaria as well as exotic animals such as leeches, ticks, poisonous snakes, and

biting insects can test the limits of tolerance. Some even claim that the side effects of anti-malaria drugs are nearly as bad as malaria itself. Certainly, these distractions can produce the same kind of errors as are produced in deserts and cold regions. Also, the alertness required to keep a lookout for pests can often be as distracting as any malaise. Natives to the jungles show remarkable adaptations to their environment, and these cases are reported in many anthropology texts.

2.3. The Overlap

People who are assigned to extreme environments but are more than just visitors—who have a designated or indefinite length of stay—can often acquire some of the adaptations of persons living permanently in extreme environments. Peace Corps volunteers on the Altiplano in Peru have adapted to the lack of oxygen from which shorter term visitors may suffer. Likewise, Federal Aviation Agency personnel in Alaska have "cold adapted" after a certain length of stay. Military personnel also report being able to adapt to cold and hot regions. A caution that needs to be applied here is that many of these individuals can have high variability in their adaptation levels depending on age, physical condition, and personal habits.

In the Cold Regions Habitability Study, Burgess Ledbetter and the author encountered a few military wives who had not been out of their houses from October to April. Some of these were suffering from the symptoms of cabin fever, but others were taking the confinement in stride. The important point is that these longer term visitors have a longer time period during which to adapt.

2.4. Isolation: The Changing Extreme

With the advent of cell phones and the ability to carry sophisticated communications equipment literally anywhere in the world, the terms of isolation are being changed. During the investigation and remedial intervention in the Biosphere 2 Study, some of the biospherians stated that they had met more people by video and phone than they had during their entire lives prior to the confinement. If this trend continues, it raises a whole new question about the quality of isolation. Will it come to the point where any worker in Antarctica, any space station astronaut, or any missionary in the field can call home at will and see the

faces of his or her spouse and children? This alternative to complete isolation cannot be overlooked if it helps to alleviate the conditions that extreme environments otherwise impose.

3. CONCLUSIONS

3.1. The Temporary Environments

What seems clear is that no extreme environment has just one variable affecting function. There are always combinations. In cold regions, it is the cold and the diurnal rhythms. In the deserts, it is the heat and the dryness. Sometimes, there are three or more variables such as cold, dryness, and diurnal rhythms. And these effects can be cumulative over time. Sleep deprivation builds over time and becomes more and more debilitating.

But extreme environments are slowly getting less extreme. Compared with the expeditions and studies done just decades ago, the communications and comforts have improved greatly. Even Antarctica can now get fresh vegetables.

Another element that has improved is the selection of personnel. After years of screening and experiments with various instruments, the selection process has virtually eliminated mental pathology from the remote stations in the world. This technology has been shared across nations, resulting in more stable environments in places such as Antarctica, the space shuttle, and other remote government stations. But is there a price to pay for this? Researchers such as Suedfeld and Palinkas have pointed out the positive benefits of isolation: the quiet, the close fellowship, and the sense of being in another world. Will this be lost in the clutter of modern communications?

Finally, will there be permanent space stations and colonies on the moon and Mars? Will people who stay there for a long time be able to return to Earth and readapt to its strong gravity? There seems to be some evidence that they will not. Stays on the space station, and even in the low-gravity colonies, might have to be limited.

3.2. The Permanent Environments

Much is being revealed about the permanent extreme environments that raises questions about the variables that affect the occupants of these environments. Migration is evident in the Altiplano, the jungles, and the deserts. Economic forces may be the most relevant of all variables, and they are sending people to places where they can earn a better living. The use of a digging tool, the *tacya,* in Peru is being replaced with machinery, making much labor surplus. The Brazilian jungle is fast disappearing, and the deserts seem to be spreading, causing a new term to be coined: desertification. These new, largely economic variables are changing not only the populations of these environments but also the actual geography itself. Political aspects enter as well, especially in places such as Tibet. But even these are changing the character of the land. Only one thing is certain: change.

See Also the Following Articles

Environmental Psychology, Overview ■ Environmental Stress

Further Reading

Ashcroft, F. (2000). *Life at the extremes: The science of survival.* Berkeley: University of California Press.

Harrison, A. (2001). *Spacefaring: The human dimension.* Berkeley: University of California Press.

Harrison, A., Clearwater, Y., & McKay, C. (Eds.). (1991). *From Antarctica to outer space.* New York: Springer-Verlag.

Nielsen, J. (2001). *Ice bound.* New York: Hyperion.

Rivolier, J., Goldsmith, R., Lugg, D., & Taylor, A. (Eds.). (1988). *Man in the Antarctic.* New York: Taylor & Francis.

Steele, G. D. (2001). Polar moods: Third quarter phenomena in the Antarctic. *Environment & Behavior*, *33*, 126–133.

Stuster, J. (1996). *Bold endeavors.* Annapolis, MD: Naval Institute Press.

Suedfeld, P., & Weiss, K. (Eds.) (2000). Antarctica. *Environment & Behavior*, *32*, 5–156.

Ward, M., Milledge, J., & West, J. (1989). *High altitude medicine and physiology.* Philadelphia: University of Pennsylvania Press.

Extroversion–Introversion

Gerald Matthews

University of Cincinnati, Cincinnati, Ohio, USA

GLOSSARY

arousal Nonspecific excitation of the brain or one of its major subsystems (e.g., the cerebral cortex).

factor analysis A family of multivariate statistical techniques that aim to explain covariance among variables in terms of a smaller number of latent factors.

interactionism A popular contemporary approach to personality that sees traits as interacting with situational factors to generate behavior.

personality disorder An enduring pattern of abnormality in personality that is typically referred to in terms of various discrete disorder types described by the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

personality trait A stable personality characteristic that describes a continuum of individual differences in behavior expressed consistently across a variety of different situations.

self-regulation The cognitive processes and behaviors that support the pursuit of personal goals within a changing external environment.

Extroversion–introversion (E–I) is a major personality dimension that is an essential part of comprehensive trait models, including the Five-Factor Model. Various questionnaires provide valid and reliable measurement of E–I. The person's level of extroversion versus introversion depends on both biological and social–cognitive factors. The mechanisms that shape the development of E–I are not well understood. On the biological side, brain systems for arousal and reward have been implicated. On the social–cognitive side, biases in various aspects of self-regulation may be important. E–I is modestly associated with many criteria related to real-life functioning, including mental health, social behaviors, and work performance. E–I should be understood within an interactionist context, such that both person and situation factors are important as determinants of behavior. Empirical studies support clinical and occupational applications based on the assessment of E–I.

1. PSYCHOMETRIC FOUNDATIONS

1.1. Origins and Early Studies

The extroversion–introversion (E–I) trait refers to a dimension of personality that contrasts qualities such as sociability, assertiveness, and enthusiasm with qualities such as social reserve, quietness, and thoughtfulness. Its place as a central element of personality has been recognized since antiquity. Jung popularized the

idea of extroverts and introverts as distinct types characterized by their tendencies to channel "psychic energy" outward and inward, respectively. However, the typological approach is no longer representative of mainstream personality research. Current research assumes that E–I is a continuous dimension or trait, such that the level of E–I is approximately normally distributed in the general population.

Dimensional models of personality traits, including E–I, are typically based on factor analysis of questionnaire data. The first questionnaire measure of E–I developed on this principle, the Inventory of Factors STDCR, was published by Guilford in 1940 and included social introversion and thinking introversion factors as two of five basic personality traits. Eysenck conducted especially influential factor-analytic research that generated a series of reliable questionnaires, beginning with the Maudsley Personality Inventory in 1956. He also pioneered the use of experimental methods for validating questionnaire assessment and investigated the role of E–I in mental illness.

1.2. Assessment of Extroversion–Introversion

A parallel approach to assessment was based on analysis of the personality-descriptive terms that appear in natural language (i.e., the lexical approach). Fiske identified five factors in personality ratings of this type, including a factor of "confident self-expression" that has been subsequently identified with extroversion. Lexical studies have converged with research based on personality questionnaires in suggesting that there are five fundamental personality traits (the "Big Five"), each of which is defined by more narrowly defined lower order traits, within a hierarchical model.

Costa and McCrae's Revised NEO Personality Inventory (NEO-PI-R) questionnaire assesses extroversion as one of the Big Five. The Five-Factor Model (FFM) has gained support from the identification of the Big Five, through factor analysis, in various questionnaires developed for other purposes. Furthermore, E–I meets four criteria listed by Costa and McCrae for a trait to be considered fundamental: E–I appears to be biologically based, it is culturally universal, it refers to important psychological concepts, and measures of E–I have good criterion and predictive validity.

There has been some controversy over the lower order traits that define extroversion as a broad trait. Eysenck distinguished sociability and impulsivity as important constituents of the trait, whereas the NEO-PI-R assesses warmth, gregariousness, assertiveness, activity, excitement seeking, and positive emotionality as primary facets. Sociability and outgoing self-confidence are universally seen as defining features of extroversion, but the place of traits such as impulsivity, activity, and dominance is still being debated. Thus, although various assessments of E–I converge, the composition of the E–I factor differs somewhat across various psychometric models. A more radical proposal was advanced by Hogan, who suggested dividing extroversion into separate factors of sociability and ambition, with the latter being seen as more predictive of job performance.

The NEO-PI-R and other questionnaire scales for E–I are internally consistent and substantially intercorrelated with one another, showing convergent validity. Measures are also stable over time, showing test–retest correlations of 0.6 to 0.7 over periods of several years. Published measures of E–I do not appear to be strongly biased by response sets such as social desirability, although such concerns should not be ignored in applied research. In addition, self-ratings of E–I converge quite well with ratings made by other persons, perhaps because E–I is one of the most salient and visible aspects of personality.

2. BIOLOGICAL BASES

2.1. Behavior and Molecular Genetics

The heritability of E–I has been investigated using twin and adoption studies. Structural equation models fitted to behavior genetic data suggest that broad heritability of E–I is approximately .50. The remaining variance is attributed to the environment, gene–environment interaction, and measurement error. Models suggest that the influence of the unshared environment, unique to each child, is considerably stronger than that of the shared environment related to the family. Molecular genetic studies are beginning to investigate specific alleles that may correlate with E–I. For example, variations at the D4 dopamine receptor gene have been found to correlate with some traits related to E–I, such as excitement seeking and positive emotionality, although the variance explained by any single allele is likely to be small.

2.2. Psychophysiological Studies

The two leading biological theories of E–I are arousal theory and reinforcement sensitivity theory (RST).

Zuckerman conducted important work on possible bio-chemical bases of E–I, but that is beyond the scope of this article. Eysenck's arousal theory identified arous-ability of a reticulo–cortical circuit controlling alertness as the key system that underpins E–I. Introverts are said to be more easily aroused, so that they regulate arousal by avoiding intense stimuli, whereas less readily aroused extroverts are prone to seeking stimulation. Numerous psychophysiological studies provide mixed support for the greater arousability of introverts, with studies of evoked potentials and phasic electrodermal responses providing the most consistent results. Replicable results are obtained only with careful control of experimental parameters.

Gray's RST links E–I to sensitivity of primarily dop-aminergic brain systems controlling reward and sensi-tivity to positive reinforcement signals. This system may be more sensitive in extroverts than in introverts, so that extroverts are more readily conditioned by reward signals than are introverts. It may also influence positive emotion. Some studies suggest that extroverts are more psychophysiologically responsive than intro-verts in rewarding situations, but the data are often inconsistent. It seems likely that multiple neural mechanisms may influence E–I. Both arousal theory and RST may have some validity, but neither fully explains the data currently available.

3. COGNITIVE–SOCIAL PERSPECTIVES

3.1. Information-Processing Studies

E–I correlates modestly with a wide variety of measures of information processing. These relationships vary with task demands. Typically, extroverts show supe-rior performance on verbal short-term memory, divided attention, speech production, and speeded motor re-sponse tasks, whereas introverts do better on sustained attention, long-term memory, and reflective problem-solving tasks. Extroverts also tend to show a lower response criterion, and favor speed over accuracy, in some task paradigms. In addition, extroverts tend to outperform introverts in arousing conditions, although this effect may reverse during evening hours.

Biological theories have difficulty in explaining the dependence of E–I effects on the information-processing demands of tasks. Instead, it seems that extroversion is associated with a complex pattern of multiple processing advantages and disadvantages relative to introversion.

Matthews proposed that the information-processing correlates of E–I have an adaptive function in preparing extroverts and introverts for the cognitive–social envi-ronments they prefer. For example, the superior verbal short-term memory and speech production skills of extroverts may support the development of conversa-tional skills that facilitate adaptation to demanding social environments.

3.2. Social Cognition

Extroverts and introverts tend to differ in various high-level aspects of cognition, including self-cognitions. Extroverts are more likely than introverts to appraise demanding situations as challenging and to cope through use of proactive problem-focused strategies. Extroverts also report greater levels of social participa-tion and affiliative needs than do introverts. Although extroversion is associated with superior social skills, the trait may also be associated with some maladaptive aspects of social cognition such as a tendency toward narcissism. In addition, extroversion has been linked to likelihood of divorce, perhaps due to impulsivity in forming and dissolving intimate relationships. Extroverts and introverts may differ in self-regulation, pursuing somewhat different goals in life supported by different styles of social cognition. However, much of the variance in social cognition is unrelated to person-ality traits and should be understood with reference to the individual's life circumstances and social learning.

4. CLINICAL APPLICATIONS

4.1. Emotional Disorders

The trait most strongly linked to anxiety and mood disorders is neuroticism, but extroversion may also be implicated. Typically, patients tend to be somewhat introverted, and introversion predicts various symptoms of emotional pathology in longitudinal data to a modest degree. Introverts are also somewhat more vulnerable to subclinical stress symptoms than are extroverts. Introversion may be especially linked to anhedonia (i.e., lack of positive affect) in depression. Many non-clinical studies suggest that introverts experience less positive affect and happiness than do extroverts, although the magnitudes of the association between E–I and affect vary considerably across studies. Some authors propose that extroverts are temperamentally disposed to experience positive affect irrespective of

circumstances, whereas others emphasize the role of situational modifiers such as social involvement.

Longitudinal studies suggest that social introversion may be a risk factor for depression, although the emotional disorder itself may raise introversion as a consequence of the patient's difficulties in social functioning. Assessment of extroversion may be useful to the clinician in understanding the patient, selecting a treatment, and anticipating the course of therapy. For example, extroverted patients may respond better to interpersonal therapy than do introverted patients.

4.2. Personality Disorders and Criminality

Traditional diagnoses of discrete personality disorders may be complemented with dimensional approaches that describe abnormal personality in terms of continuous dimensions. Dimensional studies have related several features of abnormal personality to E–I. Extroversion has been linked to excessive stimulus seeking and impulsivity, whereas introversion overlaps with social avoidance and inhibition, including schizoid personality. Thus, extremes of E–I may create a potential for personality disorders linked to these qualities. Eysenck's theory predicted that extroversion, especially in combination with neuroticism, should be related to a higher incidence of criminality and deviant behaviors because underarousal prevents effective conditioning. Although some studies have supported this prediction, the data are inconsistent. For example, persons incarcerated for long periods of time tend to show rather low levels of extroversion. It seems likely that impulsivity, rather than extroversion per se, is linked to criminality, and the more harmful aspects of impulsivity may be related more to traits such as psychoticism and low conscientiousness than to extroversion.

5. OCCUPATIONAL APPLICATIONS

5.1. Vocational Choice

Extroverts and introverts exhibit clear differences in their occupation choices. Extroverts tend to predominate in jobs that are people oriented (e.g., sales) and/or stressful (e.g., financial dealing), whereas introverts appear to be attracted to jobs that are more solitary and reflective in nature (e.g., scientific research). Similarly, studies of vocational interests show that extroverts have

stronger social and enterprising preferences. In general, E–I is only weakly related to job satisfaction, but the nature of the job seems to influence whether extroverts or introverts are better adjusted. For example, introverts show better tolerance for repetitive monotonous work than do extroverts.

5.2. Job Performance

There is an extensive literature on E–I and job performance. Often, correlations are small, as demonstrated by meta-analyses. However, such analyses may be misleading in that they neglect the likely role of the nature of the job as a factor moderating the extroversion–performance association. Research driven by a priori hypotheses has been more successful in demonstrating the occupational relevance of E–I. Extroversion relates to better performance in more jobs requiring social skills such as sales. Extroversion also relates to superior performance during training, perhaps because extroverts handle novelty better than do introverts. Conversely, introverts may perform better than extroverts in monotonous work environments. Introverts may also be less accident prone than extroverts in transportation and industrial settings, although this relationship may be mediated by impulsivity rather than by E–I per se.

6. CONCLUSIONS

E–I is a robust trait that is an essential component of any comprehensive account of human personality. There is good convergence among various questionnaire measures of the trait, although there is still debate over its constituent traits. Behavior genetic studies show that both genotype and environmental influences are important influences on phenotypic E–I. Psychophysiological and cognitive–social studies have made some progress in revealing the sources and consequences of an individual's level of E–I, but current theories of the trait are not fully satisfactory. Nevertheless, assessment of E–I may be important in a variety of applied contexts, including clinical and occupational psychology.

See Also the Following Articles

Psychometric Tests ■ Self-Control ■ Traits

Further Reading

Costa, P. T., Jr., & Widiger, A. (Eds.). (2002). *Personality disorders and the Five-Factor Model of personality* (2nd ed.). Washington, DC: American Psychological Association.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural sciences approach.* New York: Plenum.

Furnham, A., & Heaven, P. (1999). *Personality and social behaviour.* London: Edward Arnold.

Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality traits* (2nd ed.). Cambridge, UK: Cambridge University Press.

Nyborg, H. (Ed.). (1997). *The scientific study of human nature: Tribute to Hans J. Eysenck at eighty.* Amsterdam: Pergamon/Elsevier Science.

Eyewitness Identification

Colin G. Tredoux
University of Cape Town, Rosebank, South Africa

Christian A. Meissner
Florida International University, Miami, Florida, USA

Roy S. Malpass and Laura A. Zimmerman
University of Texas at El Paso, El Paso, Texas, USA

GLOSSARY

blank lineup (target absent lineup) A lineup in which all the members are known to be innocent.

cognitive interview An interview technique designed to enhance memory and communication of events through the use of guided retrieval techniques.

effective size The number of plausible members in a lineup.

encoding specificity principle If cues that were present at encoding are also present at retrieval, recall of stored information is more likely.

estimator variables Factors in eyewitness situations whose influence can only be estimated and are not under the control of the justice system.

eyewitness A person who observes an event and can provide information about that event.

lineup bias A lineup constructed in a manner that leads the suspect to stand out from the other lineup members.

mock witness A person who is not a witness to a crime but is asked to identify a perpetrator from a lineup based on another source of information, such as a verbal description.

postevent (mis)information Events that occur after a crime that influence eyewitness memories of the crime.

sequential lineup A lineup presentation procedure in which one lineup member is presented to the witness at a time.

showup The presentation of a single suspect to a witness in order to determine if that suspect is the perpetrator.

simultaneous lineup A lineup presentation procedure in which all lineup members are presented to the witness at the same time (i.e., in a line).

system variables Factors that may influence eyewitness memory and are under the control of the criminal justice system.

verbal overshadowing effect When the act of verbally describing a perpetrator decreases the witness's ability to accurately identify a perpetrator from a lineup.

An eyewitness's identification of a perpetrator is a prevalent form of incriminating evidence presented in a criminal case. However, eyewitnesses are frequently mistaken, and their errors have led to the conviction of innocent individuals. For more than a century, psychologists have sought to understand the variety of factors that may influence the accuracy of eyewitness identification. This article reviews the current state of

knowledge in the science of eyewitness identification according to the sequence of events that is likely to unfold from the time that an eyewitness makes an observation to the prosecution of the perpetrator of the crime.

1. INTRODUCTION

The testimony of eyewitnesses is an important factor in many criminal cases. Cases that hinge on eyewitness testimony can bring perpetrators of crimes to justice, set innocent people free, and enable the police and courts to carry out their prime societal responsibilities of preserving law and order. However, eyewitnesses are frequently mistaken, leading to the conviction and imprisonment of innocent people for crimes they did not commit. The extent of this problem is becoming clearer with the exoneration of innocent people through DNA evidence. According to The Innocence Project, a New York-based legal team that reviews case records of incarcerated individuals who claim their innocence, there are currently more than 140 documented DNA-exoneration cases in the United States. Mistaken eyewitness identification is generally the most frequent source of evidence used to convict an innocent defendant. Of the current database of DNA exonerations, approximately 85% are estimated to have involved faulty eyewitness identification evidence.

Acknowledgment of the dangers of eyewitness identification is not new. In England, for example, the Criminal Law Revision Committee of 1972 gave explicit recognition to this fact: "We regard mistaken identification as by far the greatest cause of actual or possible wrong convictions." Psychology and law researchers have long recognized the vagaries of eyewitness identification, conducting original empirical research from as far back as the late 19th century. This research has made a significant contribution to police and legal practice, partly because psycholegal researchers bring an empirical and scientific perspective to a legal problem, which is an approach that is quite unnatural for lawyers and jurists to take. In the current discussion of eyewitness identification, we structure our review, as far as possible, according to the sequence of events that is likely to unfold when an eyewitness makes an observation. The processes that eyewitnesses go through at each stage of the process can vary greatly, and the interests of researchers are quite different at each stage, as are the challenges they face.

1.1. Stages in the Eyewitness Process

There are three main stages of the eyewitness process: (i) perception of an event and the persons central to it, (ii) storage and assimilation of the information extracted from that experience, and (iii) remembering aspects of the experience and acting on it. Time delays between the first and third stages can vary greatly. Some information may be requested almost immediately, as when a witness at the crime scene is asked to describe an event or an offender. Other information may be requested much later, such as when a suspect is presented for identification and the witness is asked to make a complex set of judgments and decisions.

1.2. Categorization of Variables

Pertinent psychological variables differ at the various stages of the eyewitness process, and different bodies of theory and research may apply to each variable to understand them. Two important categories of variables are system variables and estimator variables, distinguished by Gary Wells in 1978. Estimator variables arise during the first stage of the eyewitness process and are those factors over which the criminal justice system has no control, whose degree and magnitude can only be estimated after the fact. By definition, these include attributes of the witness, the event, and the perpetrator, and they may involve such things as the opportunity the witness had to view the event and perpetrator, the state of mind of the witness at the time, the attentional constraints that have impinged upon the witness as a result of the demands made on the witness's attention, or any person variables (such as race or gender) that may have influenced memory performance. System variables, on the other hand, are under the control of the criminal justice system. These variables focus on the treatment of the witness, the manner in which law enforcement officers attempt to obtain information from the witness, and the interaction of person attributes with these processes. System variables occur at later stages of the process (e.g., the second and third stages), at the points where information is requested of the witness.

The demands on research vary depending on the stage of the witnessing process and the type of variable under study. For early stage estimator variables (such as the quality of memory encoding), an effective understanding requires one to simulate them in the laboratory or to study them *in situ*, in collaboration with law enforcement. The latter is rarely done. Law

enforcement generally has a very limited interest in and capability for studying these matters, although worldwide there are a few major law enforcement agencies that engage in research in this field. Many other estimator variables cannot feasibly be studied in laboratory simulations for ethical or logistical reasons, such as high levels of stress, fear, and violence. We review, in general terms, what is known about major estimator variables in this article.

System variables are quite different from estimator variables in that they are implemented in similar ways by law enforcement and researchers, and they act upon persons who have already been through the first stages of a criminal event. Thus, witnesses will naturally vary in the amount of information they have encoded about the event and the perpetrator. Witnesses in laboratory studies of system variables will often participate in crimelike events, but these are not intended as precise simulations of criminal events. Rather, they are merely ways of giving witnesses a temporal event, with some degree of complexity, and that includes information about an event and a perpetrator about which they might later be tested. Of most interest is the effect of various system variables (such as the structure of a lineup or the instructions provided to a witness) on the accuracy of information obtained from the witness.

1.3. General Research Assumptions

Generally speaking, the following assumptions underlie research in this area. First, all individuals function according to the same general set of psychological processes. For example, perception and memory do not work differently for different people in any fundamental sense, and the psychological processes that witnesses engage in do not fundamentally change in different contexts or situations. Second, information must be encoded at the time of the witnessed event for it to be reported later. If the witness does not attend to the event or to information within the event, then information about that event will not be available for subsequent recall or recognition. Any distractions of attention away from portions of an event will result in decreased encoding and therefore an absence of subsequent memory for such information. Likewise, attraction of attention to some elements will result in superior encoding of such information. Finally, the willingness of the witnesses to identify someone from a lineup involves a decision process that can be somewhat controlled by those administering the identification procedure. Although the specific factors leading to identification

may differ from person to person, the decision processes utilized are assumed to be highly consistent.

2. THE EVENT

An eyewitness is defined by an event: This is something he or she directly observes and will later be asked to provide information about. In order to study eyewitnesses, then, we need to study the kind of events to which they bear witness. However, witnessed events range from the mundane to the terrifying and are potentially infinite in their diversity. Eyewitness researchers are thus faced with the daunting task of understanding how eyewitness memory and related processes function across a wide range of possibilities. It is important to understand that eyewitness recognition and identification of the criminal offender is based (for the most part) on comparing a face (person or photo) shown to the witness with an image of a face stored in memory. In general, the lower the quality of the image stored in memory, the less likely an eyewitness is to make an accurate recognition decision. Any event or process that degrades perception (i.e., acquisition of the original image) also reduces the likelihood of a correct recognition or identification. The following factors contribute to a witness's opportunity to encode information about an event or an individual.

2.1. Time to View the Event/ Perpetrator

The duration of the witness's exposure to the offender is related to later recognition performance, such that limiting exposure time generally reduces eyewitness accuracy. This has been shown in both laboratory and archival studies. Although a range of specific times has been employed across studies, it is difficult to calibrate specific time durations to specific levels of identification accuracy, particularly given the many other factors involved. Furthermore, it is difficult to accurately assess the time a witness had to view the perpetrator *in situ*, as studies have shown that witnesses' post hoc recollection of time estimates can be very inaccurate.

2.2. Distance from the Event/ Perpetrator

The physical distance between the witness and the offender is also related to later recognition performance,

with longer distances leading to poorer rates of identification. Again, it is difficult to calibrate specific distances with specific levels of identification accuracy. In addition, witnesses can be very inaccurate at estimating the distance between themselves and the offender, so the information may not be helpful or reliable even if we were able to calibrate accuracy in a laboratory.

2.3. Visibility

Light levels (e.g., time of day) and obscured illumination (e.g., sun shining in the witness's face) have been studied by eyewitness researchers and shown to influence both perceptual and identification processes. Specifically, poor lighting and obscured illumination result in lower rates of accurate identification.

2.4. Stress, Fear, and the Presence of a Weapon

A witness's attention may be impaired or distracted if he or she focuses on the psychological stress or fear accompanying a criminal or otherwise traumatic event. Even if a witness tries to be attentive, high fear or stress may hinder the accuracy of subsequent identifications. Psychologists have often described the effects of stress or fear according to the Yerkes–Dodson law, which posits that very low or very high amounts of arousal are most likely to impede encoding and recognition. The presence of a weapon, which may be accompanied by stress or fear, has also been shown to divert a witness's attention away from the face of the offender. In general, the presence of a weapon reduces both the subsequent quality of the description provided by the witness and his or her attempts at perceptual identification of the perpetrator.

3. THE WITNESS

Individual eyewitnesses may differ along a number of dimensions or attributes, and it is natural to ask whether any of these attributes are relevant to an eyewitness' performance. Research has identified several important characteristics.

3.1. Witness Age

Although eyewitness memory remains relatively stable across a wide age range, it is less reliable for individuals in certain age groups. For example, very early in life, children demonstrate a limited ability to recall information and show lower accuracy in the identification of faces. They also have a proclivity toward choosing (or guessing) during a lineup identification task, leading to an increased likelihood of mistaken identification of an innocent person. By early teen years, their performance is generally indistinguishable, on average, from that of adults. Although children tend to recall less information when compared with adults, the overall proportion of correct information recalled does not typically differ. Older adults also become somewhat variable, but in the absence of a disease process (e.g., dementia), older adults are indistinguishable, on average, from younger adults. Sensory changes, such as cataracts or changes in the contrast or sharpness of vision, may affect an elderly witness's ability to gather and encode information about a situation or offender. Elderly adults also appear to be more susceptible to the effects of suggestive questioning or postevent misinformation than the average adult. Although children may also be susceptible to suggestion, this is presumed to be due to their unwillingness to challenge an adult's authority. The elderly are more likely to forget the sources of their information.

3.2. Alcohol and Other Drugs

There is very little research specifically on the effects of alcohol and other drugs on eyewitness memory for faces. Although research shows that alcohol somewhat impairs memory for verbal materials, its influence on face recognition or person identification tasks has varied across studies. Whereas some studies have shown no effect of alcohol on lineup identification, others have demonstrated impairment in both recall and lineup identification performance, particularly when witnesses consumed alcohol prior to viewing the event and arousal was limited. Of course, drugs that impair or disrupt perception necessarily impair encoding and memory process as well.

3.3. Witness Race, Gender, and Occupation

No reliable evidence demonstrates any general advantage or disadvantage in eyewitness memory or identification associated with the race, gender, or occupation of the eyewitness [although, it will be noted later that the race(s) or gender(s) of the eyewitness and perpetrator may interact]. However, men and women appear to differ in the type of information they recall about an

event. For example, women tend to recall more items such as clothing, whereas men focus on items such as the type of vehicle or weapon.

Law enforcement officers are commonly assumed to be superior at identifying faces and remembering the details of events. Although research has generally found that the level of experience police officers have, either in years on the job or through training, does not significantly enhance their recognition ability, it does have a positive influence on the quality of the descriptions they provide. Officers are generally able to provide more detailed accounts of an event, and they appear to be less susceptible to the effects of postevent (mis)information when compared to laypersons. Officers also provide more correct descriptive information about the perpetrator than do laypeople, without an increase in incorrect information, and tend to elaborate on action details more than laypeople. This elaboration on action details, however, must be weighed against officers' greater tendency to misperceive innocent actions as criminal in nature. Empirical studies attempting to train individuals to remember events and faces have demonstrated a similar pattern; namely, individuals' recall accuracy can be improved for details of an event, but it is difficult to improve recognition accuracy for faces.

3.4. Witness Confidence

Eyewitnesses will often claim to have great confidence in their ability to identify a perpetrator, but for more than 30 years, research examining the utility of confidence as a postdictor of accuracy has generally demonstrated a weak relationship between degree of certainty and identification accuracy. However, there are some conditions in which a stronger relationship may exist. Initial judgments made with very high confidence, for instance, have been shown to be quite diagnostic of witness accuracy.

Even though a witness's confidence may initially be related to the accuracy of the memory, that relationship can be changed dramatically by events that occur after the eyewitness makes a positive identification. If eyewitnesses are told immediately following a positive identification that they have correctly identified a suspect, not only does their degree of confidence increase but also their memory for the crime itself may change. For example, they may become more likely to report that they saw the criminal longer and under better viewing conditions than they had previously reported. In summary, although evidence exists for a relationship between confidence and accuracy under some conditions, confidence has been

shown to be quite malleable. Thus, a witness's statement of confidence, in general, may prove to be a poor indicator of identification accuracy.

4. THE PERPETRATOR

Thus far, we have considered only one of the main actors in the drama that unfolds when an eyewitness observes an event, namely the eyewitness himself or herself. The other actor is, of course, the perpetrator (or perpetrators) of the crime, and several factors have been shown to influence memory for the perpetrator.

4.1. Disguise

Disguises are frequently used during the commission of a crime, and their presence can significantly impair encoding of the perpetrator's face. Disguises may include masks, sunglasses, or anything that obscures significant portions of the face. In general, the upper portions of the face (e.g., eyes and hair) provide the most important information for later identification, and disguises that hinder the encoding of these parts of the face are more likely to prevent identification of the perpetrator.

4.2. Distinctiveness and Typicality of the Perpetrator

In general, typical faces are more difficult to distinguish from other faces in memory, resulting in a higher likelihood of false identification. In contrast, if the offender is distinctive in some way, recognition may be enhanced because the presence of unusual attributes (such as Richard Nixon's chin or Mikhail Gorbachev's forehead birth mark) can make a face easier to remember and help to distinguish it from other faces in memory. However, if the lineup identification procedure is carried out properly (as discussed later) the distinctive suspect will be placed among others who share the same distinctive attribute, requiring the witness to identify him or her without relying solely on the memorable attribute. If the witness's attention was previously drawn to the distinctive attribute, to the neglect of other attributes or away from a holistic perception of the face, then the witness's identification accuracy may be decreased.

Other facial attributes, such as attractiveness or facial "typicality," have been shown to increase the sense that a face has been seen before, resulting in an increase in

the rate of false identifications. There also appear to be a number of widely held stereotypes about general facial appearance, which influence the encoding of information about faces and may therefore influence lineup identifications by eyewitnesses. For example, faces that are stereotyped as "criminal" in appearance are more likely to be identified from a lineup.

4.3. Familiarity Due to Repeated Viewing

Not surprisingly, if the offender was previously unknown to the witness, recognition is less likely than if the offender was known to the witness. However, even when a perpetrator seems familiar, witnesses may not always be correct in their perceptions, especially if a sense of familiarity develops later in the investigative process. If the suspect was viewed committing the crime, the witness may have increased feelings of familiarity at the time of recognition and identification. However, if the suspect did not commit the crime, it is still possible that familiarity can develop after the witness views mug shots and photo spreads containing pictures of the innocent suspect. Following repeated viewings, the appearance of the suspect begins to seem familiar to the witness. Errors caused by repeated viewing have been attributed to errors in source monitoring, or source confusion. According to the source monitoring hypothesis, memory errors occur when a person attempts to identify where the memory (i.e., of the perpetrator) originated following the receipt of postevent information that leads the witness to erroneously attribute the new information (i.e., the innocent suspect) to the original witnessed event. For this reason, in-court identifications of a defendant provide little real evidence that the defendant actually committed the crime. Familiarity at this point has either been artificially strengthened (when the suspect is the offender) or created by previous viewing of the suspect's person or photo (when the suspect is not the offender).

4.4. Perpetrator Race and Gender

Although the race of the perpetrator alone generally has little influence on identification accuracy, when the race of the perpetrator and the witness are different, the interaction can impair identification accuracy. Studies of the cross-race effect or the own-race bias have shown that memory for same-race faces is generally superior to memory for faces of another, less familiar race. The effect

has been consistently demonstrated over a 35-year period and has been reliably observed with various ethnic groups (e.g., Europeans, Southern Africans, Americans, Asians, and Hispanics). The cross-race effect is of most significance to the criminal justice system when individuals mistakenly identify a suspect who is not the perpetrator.

Although female faces are generally better recognized than male faces, a similar interaction between the gender of the perpetrator and witness has also been noted. Studies of the own-gender effect have demonstrated that female participants tend to outperform male participants in remembering female faces. Curiously, though, male and female participants do not differ consistently in their ability to remember male faces.

5. OBTAINING INFORMATION FROM THE WITNESS

Eyewitness accounts of criminal events play a vital part in solving crimes and prosecuting criminals, but research has shown that very little of what witnesses actually see ever gets reported to the authorities. One reason for this limited communication has to do with the manner in which information is elicited during police interviews of witnesses. Police interview techniques are often passed on from veteran officers to rookies or learned by trial-and-error on the job. Research has shown that even police officers without formal training tend to share similar interviewing styles, and police interviews of eyewitnesses are often very loosely structured. After requesting open-ended descriptions, officers frequently interrupt witnesses by asking specific, directive questions about the crime, typically pertaining to perpetrator height and weight. These questions usually elicit brief answers from witnesses and disrupt the flow of information. This directive interview style limits the information that witnesses convey and also appears to set up a dynamic between witnesses and officers in which witnesses wait passively for officers to direct the interview. As a result, information not specifically requested by the officer may never get mentioned during the interview, despite its importance to the case.

Interview methods can seriously influence the quality of information given by witnesses in other ways, producing information that is less accurate or even fabricated. A classic example of interviewer influence on witness reports comes from the highly publicized McMartin preschool case in which accusations of ritual child sexual abuse were made against seven preschool teachers.

Videotaped interviews revealed that the children were subjected to highly suggestive and leading questions. When these methods were applied in laboratory studies, children's false allegations greatly increased. Experiences such as the McMartin case and studies of actual police interviews reveal the importance of understanding the effects of interviewer influence on eyewitness reports.

5.1. The Cognitive Interview

One interview technique developed by psychologists explicitly for the purpose of enhancing the retrieval of eyewitness memory (and limiting the detrimental effects of interviewer influence) is the cognitive interview. Developed by Geiselman and Fisher in the early 1980s, the cognitive interview consists of four main components: (i) context reinstatement, which includes mentally reinstating the environmental and personal context of the original event; (ii) instruction to "report all" information including partial information, even if it seems unimportant; (iii) recounting the event in a variety of temporal orders; and (iv) reporting the events from a variety of perspectives. A wealth of research has investigated the potential benefits of this technique, both in laboratory and field settings. A meta-analysis of these studies revealed a large increase in the number of correct details elicited by the cognitive interview and a smaller, yet significant, increase in the number of incorrect details elicited. However, the meta-analysis also indicated that accuracy rates (i.e., the proportion of correct details generated) elicited using the cognitive interview were about the same as accuracy rates using traditional interview methods (84% for the cognitive interview and 82% for standard interviews).

In subsequent research, Fisher, Geiselman, and colleagues devised the enhanced cognitive interview, a modified technique that attempted to incorporate strategies such as rapport building to manage the social dynamics of communicating effectively with a witness. These changes resulted in further increases in the amount of correct information obtained, although some studies also observed significant increases in the amount of incorrect details. Additional research is warranted. Training law enforcement officers and others to consistently employ all the aspects of the cognitive interview has proved difficult.

5.2. Guided Memory

Research shows that when various aspects of the cognitive interview are broken down, context reinstatement instructions appear to play a vital role in the effectiveness of a cognitive interview. Context reinstatement takes place when witnesses are provided with contextual cues about the to-be-remembered event. The basis for the effectiveness of context reinstatement comes from Tulving's encoding specificity principle, which maintains that retrieval cues will enhance memory when the information contained in the retrieval cues matches information contained in the original memory trace. A meta-analysis of facial recognition studies revealed that context reinstatement strategies produce some of the most substantial benefits to identification accuracy.

The guided memory technique was primarily designed to utilize context reinstatement strategies for enhancing memory. This technique involves having witnesses visualize aspects of the event, including physical features and traits of the perpetrator, along with various emotions and reactions elicited by the surrounding event. Although this method has been effective in enhancing correct recognition of perpetrators, little is known about the effects of guided memory on description quality or accuracy. Nevertheless, the guided memory technique, along with other context reinstatement strategies, has been shown to enhance the retrieval of descriptive information from witnesses and offers a promising alternative to current interview techniques.

5.3. Generating Descriptions

Archival research has shown that although descriptions provided by witnesses are frequently vague and lacking in detail, they are generally quite accurate. The descriptors most often reported by witnesses relate to action events. Besides action events, witnesses are called on to provide person descriptions, crime scene descriptions (i.e., objects and environmental details), and accounts of what was said during the crime. With regard to the perpetrator, research indicates that witnesses are much more adept at describing character traits (e.g., he looked like an accountant) and psychological attributes (e.g., he was/looked crazy, kind) than physical aspects of a face, and that such global or holistic judgments improve subsequent attempts at recognition.

5.4. Verbal Overshadowing

Although generating a description of the perpetrator is regarded as a benign activity, studies have indicated that verbally describing a target face may have a detrimental influence on subsequent accuracy at lineup identification. This phenomenon is referred to as verbal overshadowing.

Witnesses who provide a description of a target face perform worse on a subsequent identification task compared to witnesses who did not provide a description. The effect occurs regardless of whether the perpetrator is present or absent from the lineup or whether the lineup is administered sequentially or simultaneously. The negative effect is most pronounced when witnesses are encouraged to adopt a liberal response criterion and to report their memory in great detail.

5.5. Postevent (Mis)Information

Many things can happen to eyewitnesses between the time they give their report to the police and the time they subsequently attempt a lineup or in-court identification. Some intervening events can adversely (or positively) affect their identification accuracy. Eyewitnesses may read newspaper reports about the event they witnessed, they may talk to other eyewitnesses about the same event, or they may be exposed to additional information during the course of the police investigation into the crime. This is known as postevent information. Its (potentially) strongly biasing effects on eyewitness testimony have been extensively investigated and documented by Elizabeth Loftus and colleagues.

5.6. Mug Shot Collections and Composite Reproductions

Following the initial interview with the witness, police may seek to obtain a potential physical likeness of the perpetrator by having the witness attempt to recognize the perpetrator from among a library of criminal mug shots or by having the witness create a composite image based on his or her memory.

5.6.1. Searching Mug Shot Collections

Mug shot collections are maintained by many law enforcement agencies for the simple reason that criminal recidivism rates are high, and many perpetrators have been photographed by the police on a previous occasion. Researchers have investigated the influence of searching mug shot collections on later identification accuracy. This question has very important legal ramifications since eyewitnesses who have been exposed to a mug shot collection prior to making an identification of a suspect from a lineup may be more susceptible to a source monitoring error, mistaking the intervening mug shot for the perpetrator. Some studies demonstrated that mug shot

exposure can "contaminate" an eyewitness's memory, but others did not. Legally, the fact that an eyewitness has been subjected to a potentially biasing and avoidable procedure may be enough to discredit that witness.

A second, often neglected, research question concerns the cumbersome manner in which mug shot searches are conducted. Typically, large collections of photographs are stored in albums (printed or electronic), leaving the eyewitness to page through several thousand photographs in a vain attempt to "spot" the perpetrator. Some researchers have tried to improve mug shot search procedures. For example, English researchers devised a description-driven system in the late 1980s called FRAME, which appeared to work quite well but was later criticized for depending on a limited database. For instance, there were very few mug shots of old men in the database; thus, it was comparatively easy to search the database for old men, inflating the accuracy rate. A similar, description-based system has been developed in Canada, and although it appears promising, there is not enough research to draw a definitive conclusion. Researchers in Hong Kong are exploring ways of structuring mug shot searches according to the physical similarity of faces, but initial results have been disappointing.

5.6.2. Making and Utilizing Composite Portraits

Law enforcement agencies in many parts of the world rely on practical technologies to help crime witnesses reconstruct likenesses of faces. These technologies range from sketch artistry to proprietary computerized composite systems such as Identikit, Photofit, E-Fit, Mac-a-Mug, Faces, and Comp-U-Sketch. Unfortunately, these technologies have not performed well under empirical examination. In most instances, they appear to produce poor-quality composites, which are difficult to match to target faces, even when the target is in full view of the witness. There was some hope that the move to computerized composite software would result in systems containing much larger libraries of features that would be easier to search and graphically "post process" (e.g., allowing the easy addition of specific facial features such as scars or alteration of the aspect ratio of a face). There is no evidence that computerized systems (e.g., Faces and E-Fit) lead to better reconstructions than "manual" systems (e.g., Identikit and Photofit), except when the target is in full view of the witness during the reconstruction, in which case there is a clear advantage for computerized systems.

Recent improvements in computing power have produced more versatile, mathematically sophisticated composite systems. Several groups of researchers are experimenting with "eigenface" systems that potentially allow witnesses to search limitless populations of faces with the assistance of genetic algorithms. Results are promising when witnesses are allowed to reconstruct faces while they are in full view but not from memory. There is evidence that verbal descriptions may be of more assistance to eyewitnesses than visual likenesses or reconstructions.

6. THE WITNESS IDENTIFICATION

6.1. Measuring Lineup Fairness

There are two key events for an eyewitness, as far as the law is concerned. The first is the event they witnessed, and the second is the identification of the person whom they saw commits the crime. When and how the second event unfolds is just as critical as the first because this identification will constitute the evidence against the defendant. Historically, eyewitnesses were primarily asked to identify perpetrators in court, but this practice is frowned upon nowadays since it is very suggestive (indeed, the witness would look foolish if he or she pointed anywhere but at the defendant). An alternative approach is for police to arrange an encounter between the eyewitness and the accused and to hope for a spontaneous identification (i.e., a "show-up" identification). This slightly less suggestible technique still jeopardizes the liberty of innocent people.

The most widely used alternative to in-court identifications in the United States, and in many other countries, is the "lineup," in which the suspect is placed alongside a number of men (or women) of reasonably similar physical appearance and demeanor, and the witness is asked to choose the perpetrator if he or she is present in the array. Police lineups originated in English criminal law and procedure, and it is clear that the notion of "fairness" is their raison d'etre. Lineups are intended to secure an identification that can potentially incriminate the perpetrator, but one that is fair to innocent people who might be suspects. Lineups are not invariably fair in everyday practice since innocent people have been convicted on the basis of eyewitness identification from a lineup. One of the strongest strands of eyewitness research is that investigating all aspects of the fairness of lineups, from their construction to their administration. An array of methods and measures that can be applied to police lineups has been developed.

In a method known as mock witness evaluation, researchers ask people who did not see the crime event and are "blind" to the identity of the perpetrator (or innocent suspect) to try and identify the suspect in the lineup. This is usually achieved by giving mock witnesses a brief description of the perpetrator (preferably the very description the eyewitness gave to the police) and asking them to indicate the lineup member who best matches this description. If mock witnesses are able to identify the suspect at a rate greater than $1/k$ (where k is the number of lineup members), the lineup is said to be biased. The proportion of mock witnesses choosing the suspect is a measure of "lineup bias."

A second aspect of lineup fairness concerns the number of plausible foils it contains. There are records of police lineups in the United States and other countries where the suspect was the only representative of a particular race or ethnic group in the lineup. In such lineups, the suspect's identity is immediately suggested to the identifying witness. The number of plausible lineup members is referred to as the effective size and is distinguished from the nominal (or actual) size of the lineup. There are currently two measures of effective size in use: (i) a descriptive measure formulated by Malpass in 1981 and (ii) a closely related inferential measure formulated by Tredoux in 1998. For a lineup to be considered fair, it should receive favorable scores on measures of both lineup bias and lineup size.

The validity of lineup fairness measures has been evaluated in a small number of studies, with positive results. Lineup fairness is frequently assessed by researchers examining police lineups in specific legal cases. A small literature exists regarding proactive use of researchers in constructing lineups that meet fairness standards on a priori grounds. A substantial discussion of lineup fairness appeared in a special edition of the journal *Applied Cognitive Psychology* in 1999.

6.2. Lineup Instruction Effects

Some legal and criminal jurisdictions prescribe the instructions, or admonitions, that officers are to provide to eyewitnesses when presenting them with a lineup. The key question is whether these warnings have an effect on the rate of incorrect identification of innocent suspects, which is the main scourge of police lineups. Empirical studies have explored the effect of biased instructions, and several staged-crime experiments have shown that instructions that presuppose the

presence of the criminal in the lineup ("Point out the person who committed the offence") lead to much higher error rates than instructions that do not ("Note that the criminal may not be present. If the criminal is present, point him out"). The implication is that the instructions given to witnesses should be tempered quite carefully with a warning indicating the perpetrator's possible absence. Research comparing a set of clearly biased instructions with instructions usually given by the Los Angeles Police Department and a set of "more balanced" instructions showed an increase in mistaken identifications for the biased instructions but no such increase with either of the other two sets of instructions.

6.3. Alternative Strategies: Blank Lineups, Sequential Lineups, and Relative Judgments

That lineups produce high rates of false identification has been known for a long time, and researchers have puzzled over the causes. The cause was long thought to be that witnesses are led to frame their decision task as choosing the offender from among the people in the lineup, thereby selecting the best choice from among the alternatives presented. The witness enters the identification task presuming that the police would not conduct a lineup if they did not have a suspect firmly in mind and thus interprets the task as requiring him or her to identify the lineup member that the police suspect.

A number of additional strategies to reduce erroneous identifications, apart from the lineup instructions and admonitions discussed previously, have been proposed. One solution, suggested approximately 50 years ago by Glanville Williams, the famous English legal authority, is to present the witness with a blank lineup (a lineup known not to contain the offender/suspect) to "trap" people who feel they must make an identification into making a harmless one. Witnesses who choose someone from this blank lineup are clearly mistaken and are dismissed as too unreliable to complete the main identification task. Only witnesses who do not pick from the blank lineup are presented with a suspect-present lineup. Studies have employed blank lineups in staged crime experiments and discovered that witnesses who choose a member of the blank lineup are almost twice as likely to make an incorrect identification from a subsequent perpetrator-present lineup and likewise only about half as likely to correctly identify the perpetrator.

There are a number of problems introduced by such an approach. First, blank lineups are not feasible in practice since the public would quickly learn that a lineup consists of two parts, only the second of which involves the real suspect, rendering the blank lineup ineffectual. Second, it burdens law enforcement with constructing twice as many lineups as they currently construct. (The use of modern computing systems and large mug shot databases may solve this problem in the near future.) Third, law enforcement resists "throwing away" a witness who has not yet been given a chance to view the suspect on grounds that a witness may make a choice from the blank lineup and subsequently, when presented with the suspect-present lineup, enthusiastically identify the suspect, claiming that he or she was mistaken before but is now sure this is the offender. Expecting police officers to forego the potential for this "evidence" might be too much to ask.

Another structural manipulation or alteration that has attracted a great deal of attention from psychological researchers is known as the sequential lineup. Instead of presenting multiple lineups in sequence, or members of one lineup simultaneously (exposing the witness to the suspect and a number of foils at the same time), the individual lineup members are presented one at a time and the witness is required to identify the perpetrator from the members of this sequence. Witnesses are led to believe that there are more photos to be seen than they are actually shown and are instructed to decide if each photo is or is not the perpetrator. This procedure is based on an analysis of the lineup task articulated by Glanville Williams in 1963 ("The witness may ... be inclined to pick out someone, and that someone will be the one member of the parade who comes closest to his own recollection of the criminal") and psychologically interpreted by Rod Lindsay and Gary Wells in 1985. They suggested that the structure of the conventional police lineup (all members presented at once) invites witnesses to compare the members of the lineup to one another and identify the person that "best matches" the witness's memory, leading to a decision based not on the virtual correspondence of the witness's memory of the perpetrator to one of the lineup members but on the relative correspondence of the lineup members to the witness's memory, a relative judgment strategy. Presenting lineup photographs sequentially is said to inhibit eyewitnesses from relying on relative judgments and, instead, fosters absolute judgments made by comparing their memory of the perpetrator with each lineup member. Empirical studies have demonstrated that the use of sequential lineups can reduce false identifications and increase correct rejections of lineups that do not

contain the perpetrator, although a significant reduction in the rate of correct identifications in lineups that do contain the perpetrator has also been noted.

The apparent advantages of sequential lineups has made them attractive to law enforcement authorities in some jurisdictions, and they have been promoted very actively to law enforcement by some members of the eyewitness research community. Research on sequential lineups has been very important, but there is also controversy over whether the research has been extensive and careful enough to warrant promotion of sequential lineups as a replacement for traditional police lineups.

6.4. Lineups Involving Other Sensory Modes

The lineup is usually a test of visual memory, but most legal systems also recognize that additional cues to identity may reside in the voice, mannerisms, and gait of the perpetrator.

6.4.1. Earwitness Identification

The case of voice identification has attracted much attention in the research literature. The most pertinent finding has been that familiarity of the voice plays a key role in identification, and that identification of unfamiliar voices is prone to yield alarmingly high false-alarm rates. Unfamiliar voices are very difficult to recognize and especially easy to confuse. Similar to the own-race bias in facial identification, several studies have observed that voices of another, less familiar race are also more difficult to recognize. In addition, a witness's memory for the voice of the perpetrator is equally susceptible to the many estimator variables discussed previously, including opportunity to hear the voice clearly and various factors that may draw the witness's focus of attention away from the auditory stimulus.

6.4.2. Multimode Lineups

Lineups typically used in Western countries restrict the presentation of the lineup to one mode, most frequently involving visual presentation. However, cues from a single (sensory) domain may be degraded with respect to the original context (e.g., physical disguise), and it is easier to identify the perpetrator with the assistance of cues from other domains. Researchers have begun to explore the utility of multimode lineups on witness identification performance. Although there is little

published research, the idea of allowing witnesses access to cues that are not solely visual in nature is already in practice in Sweden. The lineup is formed in a room in which lineup members are allowed to sit down, smoke, communicate, and behave in general as they ordinarily would over an extended period of time. These individuals are then viewed by the witness from behind a one-way mirror, and members of the lineup are not aware when the witness is present. Unfortunately, no systematic evaluation of this strategy has been performed.

6.4.3. Live vs Photographic Lineups

Lineups currently differ within single sensory modes, such as the practice of eliciting identifications from photographic lineups as an alternative to corporeal (live person) lineups. In many areas of the United States, for example, photographic lineups are favored over corporeal lineups. In many foreign jurisdictions, however, identifications from corporeal lineups are required if eyewitness identification evidence is to be presented in court. Intuitively, corporeal lineups would seem to have an advantage: The detail to be obtained from a live inspection of the suspect must surely exceed the detail present in a photograph. However, the consensus from several studies is that there is either very little difference between the capacity of photographic and corporeal lineups to elicit identifications or no difference at all. An extensive review of the literature reported no notable difference between live, video, or photographic lineups. There is no evidence to suggest that media that embellish cues will aid identification accuracy or reduce false alarms.

7. THE PROSECUTION

After the police obtain a positive identification from the eyewitness that their suspect is indeed the perpetrator, they are ready to hand the case over for prosecution. The district attorney (or attorney general or director of public prosecutions) may decide not to prosecute, but in most jurisdictions eyewitness identification is considered strong evidence against the perpetrator and the case will proceed to court. A number of factors in the court process can influence an eyewitness's accuracy and can change the impact that an eyewitness's testimony will have at trial. For instance, the prosecution (or defense) can pose leading questions to the witness, eliciting false recollections, or an

expert can be called to testify on the vagaries of eyewitness identification. In the following sections, some of the research on two potentially important factors is summarized. First, we briefly discuss a debate that concerns the role and value of expert testimony by eyewitness researchers in legal trials. Second, we discuss research that has examined the effectiveness of jurors in assessing the reliability of eyewitness testimony at trial and whether juries benefit from expert assistance on eyewitness memory.

7.1. Expert Testimony

Many researchers have offered expert testimony to the courts in cases in which eyewitness evidence has been at issue, although this undertaking has been somewhat controversial within the discipline of psychology. A variety of arguments have surfaced on this issue, including disputes over the consistency of findings and the adequacy of current knowledge for presentation to the courts. Several surveys of researchers and experts in the eyewitness literature by Saul Kassin and colleagues have attempted to address these concerns by identifying areas of agreement. In 1989, this survey revealed that the following areas were considered "reliable" by the majority of respondents: limited exposure time, lineup instruction effects, prior expectations on the part of the witness, postevent misinformation, and the weak relationship between confidence and accuracy. An update to this survey of experts in 2001 demonstrated some additions to this listing that included confidence malleability, mug shot-induced biases, child witness suggestibility, the effects of alcohol intoxication, the cross-race effect, the weapon focus effect, forgetting over time, and lineup presentation manipulations (simultaneous vs sequential lineups). Given the developing nature of the science of eyewitness identification, it is reasonable to assume some shifts in experts' opinions regarding various phenomena. Nevertheless, Kassin and colleagues noted the remarkable consistency in ratings across the two surveys for the majority of topics covered.

7.2. Jurors' "Common Knowledge" of Problems with Eyewitness Identification

One basis for rejecting an eyewitness expert's testimony in court is the belief that factors affecting the reliability of eyewitness identification are common knowledge to the lay juror. Researchers have sought to determine exactly what laypersons know about factors affecting eyewitnesses. Three basic methodologies have been used to investigate this information: (i) surveying jury-eligible citizens as to their knowledge and beliefs, (ii) assessing jurors' ability to predict the outcome in an eyewitness identification experiment, and (iii) using mock trials to assess the influence of trial techniques (such as cross-examination of the eyewitness or the presence/absence of expert evidence on eyewitness memory).

First, survey studies have been conducted by administering questionnaires, such as the Knowledge of Eyewitness Behavior Questionnaire, in order to assess beliefs about factors that affect the accuracy and reliability of eyewitness identification. The results of such studies have demonstrated that respondents appeared insensitive to such effects as the age of the witness (young or old) and the retention interval prior to identification, and that respondents tended to believe (contrary to research findings) that training could improve identification accuracy. Second, studies have asked participants to read written summaries of identification experiments and then to "postdict" the accuracy of participant–witnesses. Results in such studies have indicated that participants overestimate accuracy rates, suggesting that individuals often believe witnesses to be much more accurate in their judgments than they truly are. Finally, research employing the "mock trial" as a method for assessing jurors' common-sense knowledge has manipulated different factors known to influence, or not to influence, identification accuracy and then assessed whether participant–jurors are sensitive to these factors in their verdicts. A common outcome is that witness confidence is given too much weight, with jurors believing confident witnesses to be more accurate, which may not necessarily be the case. In contrast, jurors failed to assign sufficient weight to other factors, such as opportunity to observe the perpetrator or the presence of a weapon, which should have been considered. Taken together, this research appears to demonstrate that lay jurors lack the requisite knowledge to appropriately evaluate eyewitness identification evidence, and that expert witnesses may be useful for providing such information at trial. Several studies have also attempted to assess the benefits of expert testimony within the mock trial paradigm. Their results have generally shown that exposure to an eyewitness expert leads jurors to more appropriately weight identification evidence based on factors known to influence the reliability of an identification.

8. CONCLUSION

In a 1978 paper that has been extraordinarily influential in the eyewitness literature, Gary Wells argued that the quintessence of eyewitness research is its application to a legal problem and advised that it is incumbent on eyewitness researchers to show the practical utility of their research. This approach has indeed yielded great rewards: One of the most important pieces in the eyewitness literature is the coconstruction of guidelines by psychologists, lawyers, and criminal justice professionals for collecting evidence from eyewitnesses, sponsored by the National Institute for Justice.

There is a cost, though, to the near-exclusive applied research focus taken by eyewitness researchers. An applied focus typically leads to research that is interested in improving practice (e.g., lowering the rate of false-positive identifications) but is less concerned with explaining how and why a particular intervention works. For example, we know that the sequential lineup returns a lower rate of false positives than the traditional police lineup, but we do not have a good answer for why this is the case. In general, eyewitness research has not appreciated the value of theoretical explanations and models. This is not surprising given the explicit applied orientation of researchers in the area, but it may now be time for researchers to redress the imbalance. A key reason for promoting such a redress is the belief that the applications that stem from applied eyewitness research could be more powerful if they were based on an explanatory model. If we understood the mechanism, for instance, that makes conservative instructions reduce the false-positive rate, then we might be able to design alternative forms of lineup that maximally reduce this rate.

We do not wish to propose that eyewitness research jettison its applied orientation: On the contrary, we wish to strengthen that orientation. The question is how this should be done. Our contention is that a return to the laboratory might, paradoxically, create a stronger applied foundation for the discipline.

See Also the Following Articles

Forensic Mental Health Assessment ■ Interrogation and Interviewing

Further Reading

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing*. Springfield, IL: Thomas.

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist*, 56, 405–416.

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law & Human Behavior*, 5, 299–309.

Malpass, R. S., & Devine, P. G. (1980). Realism and eyewitness identification research. *Law & Human Behavior*, 4, 347–358.

Malpass, R. S., & Lindsay, R. (1999). Measuring line-up fairness. *Applied Cognitive Psychology*, 13, S1–S7.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law*, 7, 3–35.

Ross, D., Read, J. D., & Toglia, M. (1994). *Adult eyewitness testimony: Current trends and developments*. London: Cambridge University Press.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139–156.

Sporer, S. L., Malpass, R. S., & Koehnken, G. (1996). *Psychological issues in eyewitness identification*. Hillsdale, NJ: Erlbaum.

Steblay, N. M., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law & Human Behavior*, 25, 459–473.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* [Booklet]. Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Tredoux, C. (1998). Statistical inference on measures of lineup fairness. *Law & Human Behavior*, 22, 217–237.

Wells, G. L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality & Social Psychology*, 36, 1546–1557.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, 54, 277–295.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law & Human Behavior*, 22, 603–647.

Eyewitness Testimony

Linda J. Levine and Elizabeth F. Loftus

University of California, Irvine, California, USA

GLOSSARY

amygdala A brain structure, located in the anterior part of the temporal lobe, that modulates the strength of emotional memories.

flashbulb memory An unusually vivid and detailed memory of a surprising and consequential event.

misinformation effect The tendency for erroneous postevent information to lead to memory distortions.

postevent information Information introduced after the experience of an event.

source amnesia The inability to recall the origin of the information that comes to mind when attempting to recall an event.

weapon focus Witnesses' tendency to focus on and remember the weapons in crime scenes, often at the expense of memory for other information.

Law enforcement officials rely heavily on eyewitnesses to identify the perpetrators of crimes and to make decisions about exactly what happened at the times of legally relevant events (e.g., what led to an accident). Eyewitness testimony also has a major impact on the verdicts returned by juries. Although eyewitnesses are often correct, and in some cases their testimony is the only evidence available, mistakes do occur and can lead to the wrong people being convicted and to the actual perpetrators of crimes remaining at large. How do these mistakes happen? Far from being a complete and objective record of past events, memory is malleable. Leading questions, differing recollections of others, and vividly imagining events at a later time can lead to memory errors and even to the creation of "memories" anew. Scientific research on the causes of eyewitness testimony errors has been used to develop recommendations for preventing errors and reducing the rates of wrongful convictions.

1. EYEWITNESS ERRORS AND WRONGFUL CONVICTIONS

It is estimated that as many as 7500 people in the United States are wrongfully convicted of serious crimes each year. One of the key factors leading to wrongful convictions is eyewitness error. Indeed, eyewitness misidentifications are responsible for more wrongful convictions than are all other causes combined. In 1996, Huff and colleagues analyzed factors that led to 205 cases of proven wrongful convictions of innocent people. Most cases resulted from a combination of errors, but slightly more than

half of all cases examined were convicted on the basis of eyewitness misidentification. In a report published the same year by the U.S. Department of Justice, Connors and colleagues analyzed 28 cases of wrongful conviction with DNA exoneration. They found that 80% of these innocent people had been convicted due to faulty eyewitness memory. These findings highlight the risks of uncritical reliance on eyewitness reports.

2. THE MALLEABILITY OF MEMORY

Memories for past events can change over time. After people first witness an event such as a crime or an accident, new information can become incorporated into memory. People may mingle memories of different events and fill in gaps in their memories based on more recent information. "Source amnesia," or the inability to recall the origin of a memory of a given event, is an important contributor to memory malleability. Over time, people forget whether information that comes to mind when attempting to recall an event originated from actual past experience, from subsequent suggestions by others, or from vividly imaging an event. Once the source of the information has been forgotten, people can confuse suggested and imagined events with the true event. Because the malleability of memory has important implications for evaluating the reliability of eyewitness testimony in court proceedings, psychologists have examined the conditions under which erroneous information is likely to be mistaken for witnessed events.

3. THE MISINFORMATION EFFECT

Research conducted by Loftus indicates that "postevent information," or information introduced after the experience of an event, can lead to errors in people's memories for witnessed events. Even subtle differences in the way in which a question is asked (e.g., changing from the indefinite article "a" to the definite article "the") can lead to distortions in memory. This was shown in a study where participants first viewed a brief film of a traffic accident. Afterward, they were asked a number of questions that were phrased in one of two forms: "Did you see the broken headlight?" or "Did you see a broken headlight?" Participants who were asked the question including the word "the" were more likely to report seeing a nonexistent broken headlight than were participants who were asked the question using the word "a".

In another study, people viewed slides of an automobile collision. Later, they were asked one of two questions: "How fast were the cars going when they hit each other?" or "How fast were the cars going when they smashed into each other?" The verbs, "collided," "bumped," and "contacted" were also used. Participants' speed estimates varied considerably depending on how the question was phrased. For example, estimates of the cars' speed averaged 40.8 miles per hour when the verb "smashed into" was used, whereas they averaged only 31.8 miles per hour when the verb "contacted" was used.

Memory errors of this type become more frequent as the gap of time increases between the occurrence of an event and the reporting of it. With increasing time, details of the actual event are forgotten or become less accessible, and the fragmentary memory tends to be filled in with information that is currently available. This information may be drawn from schematic knowledge, or knowledge of the ways in which events typically occur, from suggestions by others and even from the phrasing of a question (e.g., use of the verb "smashed" instead of "contacted"). Thus, the way in which a police officer or attorney questions a witness can influence what the witness recalls.

More broadly, scientific research shows that a person does not simply experience an event, file it away in memory, and then retrieve it whole at some later date and read off what was stored; instead, at the time of recall, past events are partially reconstructed using information from many sources. These sources include the person's original perception of the event and inferences drawn later after the fact. Over a period of time, bits of information from these sources become integrated, so that the witness becomes unable to say how he or she knows a specific detail. Only the unified reconstructed memory remains. Memories, then, are neither wholly accurate nor wholly distorted. They typically incorporate both memory traces from the experience of the actual event and postevent information.

4. CREATING NEW MEMORIES

Not only can memories be altered, they can be created anew. A simple reliable method for creating false memories for words was developed by Deese, Roediger, and McDermott and is referred to as the DRM paradigm. After studying a list of related words (e.g., rest, bed, night, snooze), people are asked whether or not particular words were on the list. They often claim to

clearly remember seeing a word that was not presented but that was closely associated with other words on the list (e.g., sleep). Hancock and colleagues' 2003 research demonstrated that these phantom words can be even more accessible in memory than words that actually appeared on the list.

Memories of complete events that never occurred also can be created in the minds of normal adults by enlisting the help of their mothers, fathers, and other older relatives. In a 1995 study, Loftus and Pickrell told participants that their relatives had provided researchers with descriptions of events that had occurred during their childhoods. In fact, three of the events had actually occurred, but one was a fabricated account of being lost in a shopping mall at 5 or 6 years of age for an extended time and ultimately being rescued by an elderly person and reunited with the family. Approximately a quarter of the participants fell sway to the suggestions and described memories, either complete or partial, of having been lost in this specific way. Using the same procedure in 2001, Heaps and Nash found that 37% of the participants fell sway to the suggestion that they were in trouble in the water—nearly drowning—and had to be rescued by a lifeguard.

Suggesting to participants that relatives had witnessed an event is a strong form of suggestion, but distortions in memory can occur even with weaker suggestions such as guided imagination. This was demonstrated in Garry and colleagues' 1996 study in which participants first rated how confident they were that a number of childhood events had happened to them before 10 years of age (e.g., "broke a window with your hand"). Later, some participants were told the following: "Imagine that it's after school and you are playing in the house. You hear a strange noise outside, so you run to the window to see what made the noise. As you are running, your feet catch on something and you trip and fall." The script went on to guide participants through a detailed scenario in which they would break the window with their hand and get cut and bloody. During the final phase of the study, they once again answered questions about their childhood experiences. The results showed that the 1-minute act of imagination led a significant minority of participants to claim that an event was more likely to have happened (relative to a control group of participants who were not asked to imagine the event), even though they had previously said the event was unlikely to have occurred.

Imagination has also been shown to influence memory for complex events such as witnessing a crime. As

with the misinformation effect, the impact of imagining an event becomes more pronounced over time. Imagining an event increases its familiarity. As both the event and the act of imagination recede into the past, people become less able to distinguish whether the source of the event's familiarity is that they imagined it or that it actually occurred.

5. CONFIDENCE AND MEMORY ACCURACY

Jurors often understand that memory is fallible and look for cues as to whether an eyewitness's testimony can be believed. One commonly used indication of memory accuracy is confidence. Juries find confident witnesses to be far more believable than those who admit that they could be mistaken. Unfortunately, the assumption that confidence is a clear sign of accuracy is unfounded. Bothwell and colleagues' 1987 review of 35 studies involving staged incidents showed that the average correlation between witness confidence and accurate identification of the perpetrator was modest (the average correlation coefficient was $r = .25$). Thus, with increasing confidence, only a slight increase in memory accuracy was found.

Further weakening its predictive power, witness confidence is highly malleable. Witnesses can become more or less confident about the accuracy of their memories for an event or individual as a result of events that happen afterward. Confidence increases when witnesses describe past events repeatedly or receive confirming feedback such as being told that another witness identified the same individual as the culprit. Confidence is more strongly predictive of recall accuracy when viewing and reporting conditions are optimal, for example, when a witness had a long time to observe a perpetrator's face and was not exposed to repeated or suggestive questioning.

6. EMOTION AND MEMORY

Witnessing a crime, and certainly being the victim of a crime, is a highly emotional event. How does intense emotion affect the accuracy and completeness of eyewitness testimony? In comparison with mundane events, memories for emotional events tend to be long-lasting, vivid, and detailed. When asked to describe how they first learned about highly emotional events (e.g., the terrorist attacks on the World Trade

Center and U.S. Pentagon on September 11, 2001), people can typically report, in considerable detail, where they were, what was happening at the time, who told them the news, and how they felt. In their classic 1977 article on flashbulb memories, Brown and Kulik argued that because of the obvious survival value, there may be some mechanism in the brain that leads to remembering biologically crucial but unexpected events with close to photographic accuracy.

More recently, animal and brain imaging studies have been conducted to investigate the biological underpinnings of the effects of emotion on memory. The results indicate that the amygdala, an almond-shaped structure located in the anterior part of the temporal lobe, modulates the strength of emotional memories stored elsewhere in the brain. The greater the level of amygdala activation, the more enduring the emotional memory. For example, in Canli and colleagues' 2000 study, participants viewed negative and neutral pictures. Functional magnetic resonance imaging (fMRI) was used to assess the level of activation in the amygdala as participants rated each picture for emotional intensity. The more emotionally intense participants found the pictures to be, the more activation was found in the amygdala. Three weeks later, participants were given a surprise recognition test with previously seen and new pictures. They were asked to indicate whether they had seen each picture before and, if so, whether it seemed familiar or whether they clearly remembered seeing it. For those pictures rated as emotionally intense, the degree of amygdala activation predicted whether the picture would be forgotten, would appear to be familiar, or would be clearly remembered.

Emotional memories, then, are more lasting and vivid than nonemotional memories, but this does not mean that they are error free. Indeed, Neisser and Harsch's 1992 research showed that inaccuracies were the norm when college students recounted how they found out about the explosion of the *Challenger* space shuttle after a delay of 2½ years. Furthermore, emotional arousal might not enhance memory for all aspects of emotional events. Several laboratory studies have shown that, in comparison with their memory for neutral events, people showed enhanced memory for information that is central to emotional events but showed poorer memory for the peripheral details. In real-world settings, a similar phenomenon has been termed "weapon focus." This refers to witnesses' tendency to focus on and remember the weapon in crime scenes where a knife or gun is present, often at the expense of memory for other information such as the culprit's face.

In summary, emotional arousal improves memory for central or gist information but can disrupt memory for details that are peripheral to a witness's current concerns (e.g., survival). Investigators are currently working to define more precisely the types of information that are "central" to people experiencing intense emotion and whether this information varies depending on the particular emotional state being experienced (e.g., fear vs anger). This issue is important because, in the case of eyewitness testimony, memory for details (e.g., a license plate number, the color of a jacket) can be telling.

What about memory for emotions themselves? Victims of crimes or accidents due to negligence may seek damages for the suffering they endured, but how accurately can people remember their own past emotional reactions? Levine and Safer's 2002 review of recent research showed that memories for past emotions, like memories for past events, can be biased by postevent information. In one study, Levine and colleagues asked people to recall how they felt when they first learned that O. J. Simpson had been acquitted of murder. Participants rated the intensity of their emotions shortly after the verdict was announced. Two months later, and then again after more than a year, they recalled how they had felt. At each time point, they also described how certain they were of Simpson's innocence or guilt. The results showed that the more people's beliefs about Simpson's guilt changed over time, the less stable their memories were for how happy or angry they felt when Simpson was acquitted. For example, those who came to believe that Simpson was guilty overestimated their feelings of anger, and underestimated their feelings of happiness, relative to their initial reports.

In another study, Safer and Levine had undergraduates recall how anxious they had felt before a major exam. Students who learned that they had done well on the exam underestimated how anxious they had felt before the exam (in comparison with a control group of students who had not yet learned their exam grades), whereas students who learned that they had done poorly overestimated how anxious they had felt before the exam. Thus, postevent information about their grades led to distortions in students' memories for their past feelings of anxiety.

Like memories of events, then, people's memories of their emotions are partially reconstructed; that is, they change subtly over time and become more consistent with current beliefs about the emotion-eliciting event. Put differently, people are constantly learning. Memories of past events are not sealed off from this subsequent learning; instead, they are informed by it.

The malleability of memory may be adaptive in some circumstances, but it makes it necessary to use caution when evaluating eyewitness testimony, whether that testimony concerns a neutral event (e.g., a traffic violation), an emotional event (e.g., a crime), or even a person's own past emotional reaction.

7. FACILITATING ACCURATE TESTIMONY

Scientists investigating the factors that lead to inaccurate memories have developed recommendations to ensure that witnesses provide as complete and accurate testimony as possible. Careful questioning of witnesses is essential for minimizing errors in eyewitness testimony. Using verbal or nonverbal cues, an interviewer can unwittingly influence witnesses toward identifying the person who the interviewer believes is the culprit and toward remembering events in the way in which the interviewer believes they occurred. Therefore, experts recommend that witnesses be interviewed by individuals who have no preconceptions concerning the facts surrounding the case.

Witnesses can be biased simply by limiting the answer options with which they are provided. Thus, interviewers are advised to phrase questions in a clear and open-ended manner that avoids suggesting that a particular response is the correct one (e.g., "They drove away in the white van, didn't they?"). Attention also should be paid to the wording of questions because the use of one term can produce a different response from the use of another, even when the two terms have similar meanings (e.g., "smashed" vs "hit").

Based on a detailed review of scientific evidence concerning eyewitness identification errors, Wells and colleagues in 1998 developed recommendations regarding the best procedures for constructing suspect lineups and photo spreads. When eyewitnesses are asked to identify a culprit from a lineup or photo spread, they tend to identify the person or photo most resembling the culprit they remember relative to the other members of the lineup. If the culprit is not present in the lineup or photo spread, this relative judgment process often leads to a false identification. Wells and colleagues recommended that eyewitnesses be informed that there is a good possibility that the perpetrator is not present in the lineup or photo spread. This makes it less likely that an innocent individual will be selected simply because he or she resembles the criminal. They also recommended that suspect lineups and photo spreads be composed of individuals who fit the basic recollection of the witness. For instance, if the perpetrator was of a certain race, having only one person of that race present in the lineup or photo spread makes that person more likely to be selected regardless of whether he or she committed the crime. Because people recall the same events differently, it may be necessary to have individual witnesses view different lineups composed of suspects who resemble each witness's description.

Witnesses' confidence in their recollections can be influenced by subsequent events and suggestions. Therefore, experts recommend that confidence be assessed and recorded at the time a perpetrator is first identified or an event is first described. Interviewers should avoid making comments to witnesses regarding their accuracy afterward because this may inflate confidence in witnesses' memories and affect the quality of testimony at later stages.

Inaccurate eyewitness testimony can lead to a double tragedy: an innocent person being punished and the actual perpetrator remaining free to victimize others. Although memory errors can never be completely eliminated, adopting identification and interview procedures that are informed by scientific evidence will minimize inaccuracies and help to safeguard the justness of the criminal justice system.

See Also the Following Articles

Child Custody ■ Child Testimony ■ Eyewitness Identification ■ Psychology and the Law, Overview

Further Reading

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, 72, 691–695.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5, 73–99.

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. *Journal of Neuroscience*, 20(RC99), 1–5.

Connors, E., Lundregan, T., Miller, N., & McEwan, T. (1996). *Convicted by juries, exonerated by science: Case studies in the use of DNA evidence to establish innocence after trial*. Alexandria, VA: National Institute of Justice.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review, 3*, 208–214.

Hancock, T. W., Hicks, J. L., Marsh, R. L., & Ritschel, L. (2003). Measuring the activation of critical lures in the Deese–Roediger–McDermott paradigm. *American Journal of Psychology*, *116*, 1–14.

Heaps, C. M., & Nash, M. (2001). Comparing recollective experience in true and false autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 920–930.

Huff, C. R., Rattner, A., & Sagarin, E. (1996). *Convicted but innocent: Wrongful conviction and public policy*. Thousand Oaks, CA: Sage.

Levine, L. J., & Burgess, S. L. (1997). Beyond general arousal: Effects of specific emotions on memory. *Social Cognition*, *15*, 157–181.

Levine, L. J., & Safer, M. A. (2002). Sources of bias in memory for emotions. *Current Directions in Psychological Science*, *11*, 169–173.

Loftus, E. F. (1996). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E. F. (2001). Imagining the past. *The Psychologist*, *14*, 584–587.

Loftus, E. F., & Pickrell, J. (1995). The formation of false memories. *Psychiatric Annals*, *25*, 720–725.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In E. Winograd, & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories*, Vol. 4: *Emory Symposia in Cognition* (pp. 9–31). Cambridge, UK: Cambridge University Press.

Reisberg, D., & Heuer, F. (2004). Remembering emotional events. In D. Reisberg, & P. Hertel (Eds.), *Memory and emotion* (pp. 3–41). New York: Oxford University Press.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, *23*, 603–647.