Sam Goldstein
Dana Princiotta
Jack A. Naglieri
Editors

Handbook of Intelligence

Evolutionary Theory, Historical Perspective, and Current Concepts

Springer

Sam Goldstein • Dana Princiotta
Jack A. Naglieri
Editors

Handbook of Intelligence

Evolutionary Theory, Historical Perspective, and Current Concepts

2015
Springer

This volume is dedicated to the thousands of children to whom I have administered IQ tests and from whom I have learned that life is far more complex than a set of scores.
To the researchers and clinicians who have come before and will follow me in the future. Science is never easy and not always popular but all we have to manage and temper our extraordinary and frequently expansive beliefs; and to Sherrie, my wonderful wife and partner.

Sam Goldstein

I am eternally grateful to my mother, Alison Helen Princiotta, for her extraordinary combination of love and wisdom; to my partner, Paul Alexander Brighton, for his unrelenting optimism and devotion; and to my nephews, Tyler James and Hunter James, for their collective enrichment of my life, coursing from New York all the way to "Oootah" (Utah).

Dana Princiotta

This book is dedicated to my wife, Kathleen Kryza, with love and admiration.

Jack A. Naglieri

"I believe in intuition and inspiration. Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution. It is, strictly speaking, a real factor in scientific research."

Albert Einstein

"Man's mind, once stretched by a new idea, never regains its original dimensions."

Oliver Wendell Holmes, Sr.

"It is not the strongest or the most intelligent who will survive but those who can best manage change."

Charles Darwin

Editor Biographies

Sam Goldstein, Ph.D. is a psychologist with areas of study in school psychology, child development, and neuropsychology. He is licensed as a psychologist and certified as a developmental disabilities evaluator in the State of Utah. Dr. Goldstein is a Fellow in the National Academy of Neuropsychology and American Academy of Cerebral Palsy and Developmental Medicine. He is a Board Certified Pediatric Neuropsychologist. Dr. Goldstein is an Adjunct Assistant Professor, Department of Psychiatry, University of Utah School of Medicine. Since 1980, Dr. Goldstein has worked in a private practice setting as the Director of a multidisciplinary team, providing evaluation, case management, and treatment services for children and adults with histories of neurological disease and trauma, autism, learning disability, adjustment difficulties, and attention deficit hyperactivity disorder. Dr. Goldstein is on staff at the University Neuropsychiatric Institute. He has served as a member of the Children's Hospital Craniofacial Team. He has also been a member of the Developmental Disabilities Clinic in the Department of Psychiatry at the University of Utah Medical School.

Dr. Goldstein has authored, coauthored, or edited 50 clinical and trade publications, including 18 textbooks dealing with managing children's behavior in the classroom, genetics, autism, attention disorders, resilience, and adult learning disabilities. With Barbara Ingersoll, Ph.D., he has coauthored texts dealing with controversial treatments for children's learning and attention problems and childhood depression. With Anne Teeter Ellison, he has authored *Clinician's Guide to Adult ADHD: Assessment and Intervention*. With Nancy Mather, Ph.D., he has completed three texts for teachers and parents concerning behavioral and educational issues. With Michael Goldstein, M.D., he has completed two texts on attention deficit hyperactivity disorder. He has edited three texts with Cecil Reynolds, Ph.D., on neurodevelopmental and genetic disorders in children. With Robert Brooks, Ph.D., he has authored 12 texts including *Handbook of Resilience in Children,* first and second editions; *Understanding and Managing Children's Classroom Behavior,* Second Edition; *Raising Resilient Children*; *Nurturing Resilience in Our Children*; *Seven Steps to Help Children Worry Less*; *Seven Steps to Anger Management*; *The Power of Resilience*; *Raising a Self-Disciplined Child*; *and Raising Resilient Children with Autism Spectrum Disorders*. With Jack Naglieri, he has authored a number of texts on autism, assessment of intelligence, and executive functioning. He has coauthored a parent training program and is currently completing a number of additional texts on resilience, intelligence, and genetics. Dr. Goldstein is the Editor-in-Chief of the *Journal of Attention Disorders* and serves on six editorial boards. He is also the Coeditor of the *Encyclopedia of Child Development and Behavior*.

With Jack Naglieri, Ph.D., Dr. Goldstein is the coauthor of the *Autism Spectrum Rating Scales*, *Comprehensive Executive Functioning Inventory*, *Rating Scales of Impairment*, and with Dr. Naglieri and J. P. Das the *Cognitive Assessment System*, Second Edition.

Dr. Goldstein, a knowledgeable and entertaining speaker, has lectured extensively on a national and international basis to thousands of professionals and parents concerning attention disorders in children, resilience, depression, adjustment and developmental impairments, autism, and assessment of brain dysfunction.

Dana Princiotta, Ph.D. is a licensed psychologist in the state of Arizona. She completed postdoctoral study at the Neurology, Learning, and Behavior Center in Salt Lake City, Utah. In addition to this text, she has coauthored five book chapters and a peer-reviewed article.

Jack A. Naglieri, Ph.D. is a Research Professor at the Curry School of Education at the University of Virginia, Senior Research Scientist at the Devereux Center for Resilient Children, and Emeritus Professor of Psychology at George Mason University. He is a Fellow of APA Divisions 15 and 16, recipient of the 2001 Senior Scientist Award for APA Division 16, and the 2011 Italian American Psychology Assembly Award for Distinguished Contributions to Psychology. He is a Diplomate in Assessment Psychology, has earned a license as a School Psychologist in Virginia and Ohio, and earned school psychology certifications in New York, Georgia, Arizona, and Ohio. Dr. Naglieri has focused his professional efforts on theoretical and psychometric issues concerning intelligence, cognitive interventions, diagnosis of learning and emotional disorders, and theoretical and measurement issues pertaining to protective factors related to resilience.

Dr. Naglieri is the author or coauthor of more than 300 scholarly papers, books, and tests. His scholarly research includes investigations related to exceptionalities such as mental retardation, specific learning disabilities, giftedness, and attention deficit disorder; psychometric studies of tests such as the Wechsler Scales of Intelligence, Cognitive Assessment System, and the Kaufman Assessment Battery for Children; examination of race, gender, and ethnic differences in cognitive processing; fair assessment using nonverbal and neurocognitive processing tests; identification of gifted minorities, IDEA, and identification of specific learning disabilities; and cognitively based mathematics interventions. He has authored various books, including *Essentials of CAS Assessment* (Naglieri, 1999), and coauthored other books including *Assessment of Cognitive Processes: The PASS Theory of Intelligence* (Das, Naglieri, and Kirby, 1994); *Helping Children Learn: Intervention Handouts for Use at School and Home*, Second Edition (Naglieri and Pickering, 2010); *Essentials of WNV Assessment* (Brunnert, Naglieri, and Hardy-Braz, 2009); and *Helping All Gifted Children Learn: A Teacher's Guide to Using the NNAT2* (Naglieri, Brulles, and Lansdowne, 2009). Dr. Naglieri has also coedited books such as *Handbook of Assessment Psychology* (Graham and Naglieri, 2002), *Assessment of Autism Spectrum Disorders* (Goldstein, Naglieri, and Ozonoff, 2009), *Assessing Impairment: From Theory to Practice* (Goldstein and Naglieri, 2009), and *A Practitioner's Guide to Assessment of Intelligence and Achievement* (Naglieri and Goldstein, 2009).

Dr. Naglieri's scholarly efforts also include development and publication of tests and rating scales. He began this work in the mid-1980s with the publication of the *Matrix Analogies Tests* (Naglieri, 1985) and the *Draw-A-Person Quantitative Scoring System* (Naglieri, 1988) and *DAP: Screening Procedure for Emotional Disturbance* (Naglieri, McNeish, and Bardos, 1991). He published the *Devereux Behavior Rating Scale—School Form* (Naglieri, LeBuffe, and Pfeiffer, 1993), *Devereux Scales of Mental Disorders* (Naglieri, LeBuffe, and Pfeiffer, 1994), and the *Devereux Early Childhood Assessments* (LeBuffe and Naglieri, 2003). In 1997, he published the *General Ability Scale for Adults* (Naglieri and Bardos, 1997), *Cognitive Assessment System* (Naglieri and Das, 1997), and *Naglieri Nonverbal Ability Test—Multilevel Form* (Naglieri, 1997). He published the *Naglieri Nonverbal Ability Test*, Second Edition (Naglieri, 2008); the *Wechsler Nonverbal Scale of Ability* (Wechsler and Naglieri, 2008); and the *Devereux Elementary Student Strength Assessment* (LeBuffe, Shapiro, and Naglieri, 2009). Most recently, he published the *Cognitive Assessment System*, Second Edition (Naglieri, Das, and Goldstein, 2013); *Comprehensive Executive Function Scale* (Naglieri and Goldstein, 2013); and the *Autism Spectrum Rating Scale* (2010). For more information see: www.jacknaglieri.com.

Contents

Contributors

Maria J. Avitia University of Connecticut Storrs, Willington, CT, USA

Lesley Baker Department of Psychology, University of Victoria, Victoria, BC, Canada

Mark Benisz, Psy.D. Kiryas Joel School District, Monroe, NY, USA

Lisa S. Blackwell, Ph.D. Mindset Works, Walnut, CA, USA

Douglas Blancero JP Associates, Red Hook, NY, USA

Richard E. Boyatzis Case Western Reserve University, Cleveland Heights, OH, USA

Justin M. Deonarine Multi-Health Systems, Toronto, ON, Canada

Emily C. Duggan, M.Sc. Department of Psychology, University of Victoria, Victoria, BC, Canada

Ron Dumont, Ed.D. Department of Psychology, Fairleigh Dickinson University, Teaneck, NJ, USA

Amber Esping College of Education, Texas Christian University, Fort Worth, TX, USA

Dawn P. Flanagan, Ph.D. Department of Psychology, St. John's University, Queens, NY, USA

Christine Fleener, M.A. Department of Comparative Human Development, Institute for Mind and Biology, University of Chicago, Chicago, IL, USA

James R. Flynn POLS Department, University of Otago, Dunedin, New Zealand

Mauricio A. Garcia-Barrera, Ph.D. Department of Psychology, University of Victoria, Victoria, BC, Canada

James Gaskin Marriott School of Management, Brigham Young University, Provo, UT, USA

David C. Geary, Ph.D. Department of Psychological Sciences, Interdisciplinary Neuroscience Program, University of Missouri, Columbia, MO, USA

Sam Goldstein, Ph.D. University of Utah, Salt Lake City, UT, USA

John D. Greenwood, Ph.D. CUNY Graduate Center, New York, NY, USA

Belén Guerra-Carrillo, B.A. University of California, Berkeley, CA, USA

Michel A. Hofman, Ph.D. Netherlands Institute for Neuroscience, Amsterdam, The Netherlands

Arthur MacNeill Horton Jr., Ed.D., A.B.P.P., A.B.N. Neuropsychology section, Psych Associates of Maryland, LLC, Towson, MD, USA

Christopher Jones Longwood University, Hull Building, Farmville, VA, USA

James C. Kaufman Department of Educational Psychology, Neag School of Education, University of Connecticut, Storrs, CT, USA

Philip Lieberman George Hazard Crooker University, Providence, RI, USA

Dario Maestripieri, Ph.D. Comparative Human Development, Evolutionary Biology, and Neurobiology, The University of Chicago, Chicago, IL, USA

Tara Mandalaywala, Ph.D. Department of Psychology, New York University, New York, NY, USA

Department of Comparative Human Development, Institute for Mind and Biology, University of Chicago, Chicago, IL, USA

Nancy Mather, Ph.D. University of Arizona, Tucson, AZ, USA

Elias Mpofu Faculty of Health Sciences, University of Sydney, Lidcombe, NSW, Australia

Ulrich Müller Department of Psychology, University of Victoria, Victoria, BC, Canada

Jack A. Naglieri, Ph.D. Department of Virginia, Centerville, VA, USA

Thomas Oakland University of Florida, Gainesville, FL, USA

Samuel O. Ortiz St. John's University, Jamaica, NY, USA

Tulio M. Otero, Ph.D. The Chicago School of Professional Psychology, Chicago, IL, USA

Jonathan A. Plucker Department of Educational Psychology, Univeristy of Connecticut, Neag School of Education, Storrs, CT, USA

Dana Princiotta, Ph.D. Tucson, AZ, USA

Cecil R. Reynolds, Ph.D. Texas A&M University, Austin, TX, USA

Jordan Rigby, Psy.D. Neurology, Learning and Behavior Center, Salt Lake City, UT, USA

Sylvia Rodriguez Mindset Works, Walnut, CA, USA

Donald H. Saklofske Department of Psychology, University of Western Ontario, Social Science Centre (7th floor), London, ON, Canada

Deborah Schneider Doctoral Candidate and Department Fellow in Special Education, University of Arizona, Tucson, AZ, USA

W. Joel Schneider Department of Psychology, Illinois State University, Normal, IL, USA

Steven J. Stein, Ph.D. Multi-Health Systems, Toronto, ON, Canada

Robert J. Sternberg Department of Human Development, B44 MVR, Cornell University, Ithaca, NY, USA

Tarmo Strenze, Ph.D. Institute of Social Studies, University of Tartu, Estonia, Tartu, Estonia

Lisa A. Suzuki, Ph.D. Department of Applied Psychology, New York University, New York, NY, USA

Peggy L. Tarpley Department of Education and Special Education, Longwood University, Farmville, VA, USA

Kayla Ten Eycke Behavioural Research Unit, Alberta Children's Hospital, University of Calgary, Calgary, AB, Canada

Fons J.R. van de Vijver Department of Culture Studies, Tilburg University, Tilburg, The Netherlands

Hongguo Wei Case Western Reserve University, Cleveland, OH, USA

John O. Willis, Ed.D. Department of Education, Rivier University, Peterborough, NH, USA

Thomas R. Zentall Department of Psychology, University of Kentucky, Lexington, KY, USA

Part I

Introduction

1 The Evolution of Intelligence

Sam Goldstein

Viewed narrowly, there seem to be almost as many definitions of intelligence as there are experts asked to define it.

R.J. Sternberg

In the last 3,000 years of the written word, intelligence has been defined in multiple ways, including the capacity for abstract thought, understanding, communication, planning, learning, reasoning, and, most importantly, problem solving. Though most widely studied within the human species, the concept and behaviors related to intelligence have been observed and studied in animals and even plants. It is still the case, however, that there is no scholarly consensus as to what exactly defines intelligence. It is very clearly a construct that has resonated in the minds of philosophers, scientists, psychologists, and educators. It is a concept that has intrigued the general public and been used as the defining criteria for Mensa, an organization touting itself as allowing membership to only the "most intelligent." Certainly in the last 100 years, a psychometric approach has been used to define the concept in humans and in doing so offer a means of comparison between individuals.

That being said, how can intelligence be conceptualized within an evolutionary framework? What drove all species to develop abilities and acquire knowledge to enhance their survival? Has the evolution of intelligence been driven by the environmental pressures inherent in the ecosystem of any species? In this volume, we, along with our contributors, will explore and attempt to answer these questions.

The Vocabulary of Intelligence

The word intelligence derives from the Latin verb *intelligere*. This verb finds its roots in *interlegere*, meaning to "pick out" or discern (Traupman 2007). A form of this verb, intellectus, was the medieval technical term for someone with a good understanding as well as a translation for the Greek philosophical term nous. Nous, however, was strongly linked to the metaphysical, cosmological theories of teleological scholasticism, including theories of the immortality of the soul and the concept of the active intellect. Its entire approach to the study of nature, however, was rejected by modern philosophers, including Francis Bacon, Thomas Hobbes, John Locke, and David Hume, all of whom preferred the word "understanding" in their English philosophical works. Hobbes, in his work *Latin De Corpore*, used the term "intellectus intelligit" or "the understanding understandeth" as a typical example of a logical absurdity. The term intelligence, therefore, has become less common in the English language philosophy but has been widely adopted in contemporary psychology, absent the scholastic theories which it once implied.

S. Goldstein (✉)
University of Utah, 230 South 500 East, Suite 100,
Salt Lake City, UT 84102, USA
e-mail: info@samgoldstein.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_1,

In 1994, an editorial statement in the *Wall Street Journal* by 52 researchers defined intelligence as:

> A very general mental capacity that among other things involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill or test taking smarts. Rather it reflects a broader and deeper capability for comprehending our surroundings, catching on, making sense of things, or figuring out what to do. (Gottfredson 1997 p. 20)

A report published by the Board of Scientific Affairs at the American Psychological Association in 1996 titled *Intelligence: Knowns and Unknowns* defined intelligence as (Neisser et al. 1996):

> Individuals differ from one another in their ability to understand complex ideas to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: A given person's intellectual performance will vary on different occasions, in different domains as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena. Although considerable clarity has been achieved in some areas, no such conceptualization has yet answered all the important questions and non-commands universal assent. Indeed, when two dozen prominent theorists are recently asked to define intelligence they gave two dozen somewhat different definitions. (p. 77)

It seems to us that in intelligence there is a fundamental faculty, the alteration or the lack of which is of utmost importance in practical life. This faculty of judgment is otherwise called good sense, practical sense, initiative, or the faculty of adapting one's self to circumstances (Binet and Simon 1905). Many of the most influential psychologists in the field of intelligence have also weighed in on the concept of intelligence, defining intelligence as:

> The aggregate or global capacity of the individual to act purposely, to think rationally and deal effectively with his environment. (Wechsler 1944, p. 3)
>
> Intelligence is not a single, unitary ability, but rather a composite of several functions. (Anastasi 1992, p. 613)
>
> We shall use the term intelligence to mean the ability of an organism to solve new problems…. (Bingham 1937)
>
> …the ability to plan and structure one's behavior with an end in view. (Das 1984, p. 35)
>
> An intelligence is the ability to solve problems or to create products that are valued within one or more cultural settings. (Gardner 1993)
>
> … performing an operation on a specific type of content to produce a particular product. (Guilford 1967)
>
> Intelligence is the ability to learn, exercise judgment and be imaginative.
>
> Intelligence is a general factor that runs through all types of performance. (Jensen 1998)
>
> Intelligence is assimilation to the extent that it incorporates all the given data of experience within its framework. (Piaget 1963)
>
> … ability to carry on abstract thinking. (Terman 1922)
>
> The term intelligence designates a complexly interrelated assemblage of functions, no one of which is completely or accurately known in man…. (Yerkes and Yerkes 1929)
>
> The ability to deal with cognitive complexity. (Gottfredson 1998, p. 25)
>
> … goal directed adaptive behavior. (Sternberg 1984, p. 3)
>
> Unique propensity of human beings to change or modify the structure of their cognitive functioning to adapt to the change and demands of a life situation. (Feurestein et al. 1979/2002)
>
> … the mind is not a complex network of general capabilities such as observation, attention, memory, judgment and so forth, but a set of specific capabilities, each of which is to some extent independent of the others and has developed independently. (Vygotsky 1978)

Most certainly, the evolution of intelligence as a force driven by need over millions of years across multiple species has led to often fascinating ideas and conceptual processes.

The time line of human evolution spans approximately seven million years (Berstrom and Dugatkin 2011) from the separation of the *Pan* genus until the emergence of behavioral modernity 50,000 years ago. The first three million years of this time line concern *Sahelanthropus*. The next two million years concern *Australopithecus*, while the final two million years span the actual history of human (Paleolithic) traits considered to reflect human intelligence, including empathy, theory of mind, mourning, and ritual, and the use of symbols of tools is apparent in great apes as well as other species although in lesser sophistication than in humans (Matsuzawa 2001). However, the great apes demonstrate considerable ability for cognition and empathy. Chimpanzees are capable of making tools, using them to acquire food and for social displays. They are strategic hunters working cooperatively and organized by seniority. They are status conscious, manipulative, and capable of deception. They can learn to use symbols and understand human language, including some level of relational syntax and concepts of number and numerical sequence (Boysen and Berntson 1989; Hirata 2009).

It is thought that ten million years ago when the Earth's climate entered a cooler and dryer phase leading to the ice age 2.6 million years ago, one consequence was that the North African tropical forest began to retreat, being replaced first by open grasslands and eventually by desert. It is believed this forced tree-dwelling animals to adapt to this new environment. When the environment changed from continuous forest to patches of forest separated by expanses of grassland, some primates adapted to a partly or fully ground-dwelling life.

Early hominid species may have adapted to this challenge by evolving to pedalism or walking on their hind legs. This gave their eyes greater elevation and the ability to see approaching danger from further off. It also freed the arms from the task of walking and made the hands available for tasks such as gathering food. At some point, bipedal primates developed handedness, giving them the ability to pick up sticks, bones, and stones and use them as weapons or as tools for tasks such as killing smaller animals, cracking nuts, or cutting up carcasses. These primates developed the use of primitive technology. Bipedal tool-using primates from a Hominina subtribe of which the earliest species such as *Sahelanthropus tchadensis* date to about seven to five million years ago. From that point forward, the hominid brain began to develop rapidly in both size and differentiation of function. These early patterns of behavior very clearly reflect an effort to problem solve and adapt to the environment and thus likely the origins of modern intelligence. These early species needed to learn to plan, attend to relevant detail, simultaneously process multiple sources of information, and sequence their activities in order to enhance survival.

Homo habilis appeared in East Africa 2.4 million years ago. The first known human species, *Homo habilis*, was also the first known species to make stone tools. The use of such tools conferred a crucial evolutionary advantage that required a larger, more sophisticated brain to coordinate the fine hand movements required for these tasks. The evolution of a larger brain created a problem for early humans, however. A larger brain requires a larger skull, and thus, females have to develop a wider birth canal for a newborn's larger skull to pass through. However, if the female's birth canal grew too wide, her pelvis would be so wide that she would lose the ability to run, still a necessary skill in the dangerous world of two million years ago. Evolution's solution was to give birth at an early stage of fetal development before the skull was too large to pass through the birth canal and have the plates of the skull unconnected. These adaptations enabled the human brain to continue to grow, but it imposed a new discipline to care for helpless infants for long periods of time and force humans to become less mobile. Unlike salmon, snakes, bears, or primates, the human species appeared to abandon all efforts at self-preservation (thick skin or hair) in exchange for developing big brains. This greatly expanded childhood and required humans to increasingly stay in one place for longer periods of time so females could care for infants while males hunted and fought with other bands competing for food sources. As a result, humans became even more dependent on tool making to compete with other animals and other humans, relying less on body size and strength for survival. About 200,000 years ago, Europe and the Middle East were colonized by Neanderthals. They were extinct by 20,000 following the appearance of modern humans in the region some 40,000 years ago (Klein 2003). In this same period of time, Neanderthals inhabited Europe, and *Homo sapiens* first appeared in East Africa. It is unclear to what extent these early modern humans had developed language, music, or religion, but they spread through Africa over the following 50,000 years or so. It is thought that their ability to harness language to communicate and problem solve as well as the capacity to think one thing but say or do something else facilitated their spread and offered an advantage over other hominid species. Rapidly increasing sophistication and tool making and behaviors apparent by 80,000 years ago in the migration out of Africa followed toward the very end of the Middle Paleolithic some 60,000 years ago. Fully modern behaviors, including art, music, self-ornamentation, trade, burial rights, etc., are evident by 30,000 years ago. The oldest unequivocal example of prehistoric art dates to this period. Thus, it would appear that human intelligence did not evolve as a means of just solving ecological problems but as a means of surviving and reproducing in large and complex social groups. Some of the behaviors associated with living in large groups include reciprocal altruism, deception, and coalition formation. These group dynamics relate to the theory of mind or the ability to understand the thoughts and emotions of others. As hominids started living in large groups, selection must have favored effective problem solving leading to greater intelligence. Yet not all theorists agree. Geoffrey Miller argues that human intelligence is unnecessarily sophisticated for the needs of hunter-gatherers to survive. He argues that intelligent behaviors such as language, music, and art were of no utilitarian value to the survival of ancient hominids. Rather, intelligence may have been a fitness indicator. Hominids would have selected for greater intelligence as a proxy for healthy genes, and a positive feedback loop of sexual selection would have led to the evolution of human intelligence in a relatively short period of time. Such a theory must also explain why both genders are intelligent. In many species, only males have impressive ornaments and show-off behaviors. However, it is thought that sexual selection may act on males and females even in species that are at least partially monogamous. Sexual selection has clearly occurred for other female-specific human traits, for example, breasts and buttocks. Sexual selection for intelligence and judgment can act as indicators of success such as highly visible displays of wealth. Finally, if it is possible for females to successfully judge male intelligence, they must be intelligent themselves.

As advancement, survival, and reproduction within an increasing complex social structure favored even more advanced social skills, conceptual ability, reasoning, and problem solving, it is not difficult to appreciate the rapid growth of intelligence in our species. Since competition shifted bit by bit from controlling nature to influencing other humans, it becomes of relevance to outmaneuver other members of the group, seeking leadership or acceptance by means of more advanced social skills and problem-solving abilities.

There is also no doubt that higher intellectual functioning develops better in an environment with adequate nutrition. Deficiency in iron, zinc, protein, iodine, B vitamins, omega 3 fatty acids, magnesium, and other nutrients either in the mother during pregnancy or in the child during development results in impaired development (Frensham et al. 2012). The contribution of nutritional factors, including shifting from agrarianism to meat-eating diets, is thought to also have made a significant contribution to the rapid growth in intelligence and human species over the last 50,000 years.

About This Book

Tracing the evolutionary and cultural roots of intelligence is neither easy nor straightforward. As far as we are aware, these two topics have never been extensively presented simultaneously in a single volume. Following the introduction the book begins with seven background chapters. Our contributors explore the concept of intelligence in non-primates and nonhuman primates. As author Philip Lieberman points out, our species took a fast track in developing intelligence as we concomitantly developed language. Author Michael Hoffman emphasizes our knowledge of the evolution of the human brain and the conceptual development of intelligence. The final three chapters of the Background section explore philosophers beginning with Plato and Pascal through Darwin, Charcot, and Galton. The Background section ends with an important discussion by David Geary on the role social competition has likely played in pushing the evolution of fluid intelligence. The third section contains eleven chapters exploring and differentiating historical roots of intelligence beginning with early scientific researchers such as James, Wundt, and Cattell, discussing the work of Piaget and Binet and progressing through the pivotal work of A. R. Luria, setting the stage for current conceptualization of multiple intelligences, parallel types of intelligence, and intelligence as a malleable, creatively driven concept. The next section contains five chapters presenting the long and fascinating history between theories of intelligence and intelligence tests. Chapter 20 by Jack A. Naglieri explores the past 100 years of intelligence testing and offers a glimpse into the future. The next three chapters discuss applications of intellectual theory as it relates to the economic structure of our society, executive functioning, and educational achievement. Our book concludes with a seminal chapter by James Flynn and our concluding comments.

References

Anastasi, A. (1992). What counselors should know about the use and interpretation of psychological tests. *Journal of Counseling and Developments, 70*, 610–615.

Berstrom, C. T., & Dugatkin, L. A. (2011). *Evolution*. New York: W.W. Norton and Company.

Binet, A., & Simon, T. (1905). Methods nouvelles por le diagnostic du niveal intellectual des anormaux. *Nal' Annë Psychologique, 11*, 191–244.

Bingham, W. V. (1937). *Aptitudes and aptitude testing*. New York: Harper and Brothers.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (Pantroglodytes). *Journal of Comparative Psychology, 103*, 23–31.

Das, J. P. (1984). Aspects of planning. In J. R. Kirby (Ed.), *Cognitive strategies and educational performance* (pp. 35–50). Orlando: Academic.

Feurestein, R., Feurerstein, S., Falik, L., & Rand, Y. (1979/2002). Dynamic assessments of cognitive modifiability. ICELP Press: Jerusalem.

Frensham, L. J., Bryan, J., & Parletta, N. (2012). Influences of micronutrient and omega-3 fatty acid supplementation on cognition, learning, and behavior: Methodological considerations and implications for children and adolescents in developed societies. *Nutrition Reviews, 70*, 594–610.

Gardner, H. (1993). *Frames of mind: Theory of multiple intelligences*. California: Fontana Press.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence, 24*(1), 13–23.

Gottfredson, L. S. (1998). The general intelligence factor. *Scientific American Presents, 9*, 24–29.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Hirata, Y. (2009). Factors affecting the perception of Japanese moraic rhythm by second language learners. *Journal of the Phonetic Society of Japan, 13*(3), 33–43.

Jensen, A. R. (1998). *The g factor: The science of mental ability* (Human evolution, behavior and intelligence). Westport: Praeger.

Klein, R. G. (2003). Whither the Neanderthals. *Science, 299*, 1525–1527.

Matsuzawa, T. (2001). Primate foundations of human intelligence: A view of tool use in non-human primates and fossil hominids. In M. Matsuzawa (Ed.), *Primate origins of human cognition and behavior*. New York: Springer.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, W., Brody, N., et al. (1996). Intelligence: Knowns and unknows. *American Psychologist, 51*, 77–101.

Piaget, J. (1963). *The psychology of intelligence*. New York: Routledge.

Sternberg, R. J. (1984). Toward a triarchic theory of human intelligence. *Behavioral and Brain Sciences, 7*, 269–287.

Terman, L. M. (1922). A new approach to the study of genius. *Psychological Review, 29*(4), 310–318.

Traupman, J. (2007). *The Bantam new college Latin and English dictionary* (3rd revised ed.). New York: Bantam.

Vygotsky, L. S. (1978). Interaction between learning and development. In L. Vygotsky (Ed.), *Mind and society* (pp. 79–91). Cambridge: Harvard University Press.

Wechsler, D. (1944). *Measurement of adult intelligence*. Baltimore: The Williams and Wilkins Company.

Yerkes, R. M., & Yerkes, A. W. (1929). *The great apes: A study of anthropoid life*. New Haven: Yale University Press.

Part II

Background

2 Intelligence in Nonprimates

Thomas R. Zentall

Humans tend to have an anthropocentric view of intelligence that views them at the top and quite often animals that look like us close behind. Although the notion of an evolutionary scale with humans at the top is popularly held, it is also self-serving. We tend to overvalue our problem-solving ability, our capacity to modify our environment, and our ability to communicate with each other. Conversely, we tend to undervalue the exceptional sensory skills of other animals, for example, the tracking and drug-detecting ability of dogs; the navigational abilities of homing pigeons, whales, and monarch butterflies; and the ability of birds of prey to detect the minute movement of a small animal on the ground far below them. The role of our intelligence in the domination of our species over others seems obvious, but in the broader perspective of evolutionary success, as measured by the number of surviving members of a species, intelligence, as a general characteristic, correlates somewhat negatively with most measures of evolutionary success. Consider the relatively small numbers of our closest relatives, the great apes, compared with the large numbers of physiologically simpler insects, bacteria, and viruses. And it is estimated that if a massive disaster were to occur (e.g., if the Earth were hit by a large asteroid or suffered a self-inflicted nuclear disaster), many simpler organisms would likely survive much better than large intelligent animals like us.

From a purely biological perspective, the ideal survival machine is a simple, one-celled, organism (e.g., the amoeba) that has survived as a species in one of two ways. Either it has needed to undergo little change in morphology or behavior for millions of years because it exists in a remarkably stable (predictable) environment, in which case there has been little need for change, or if its environment does change, it relies on natural selection by means of very rapid reproduction and mutation (e.g., bacteria and viruses). This ability to reproduce quickly and often, ensures the survival of many of these organisms (albeit not necessarily in the same form) even in the event of a major catastrophe. Many other organisms whose rate of reproduction has not been able to keep up with relatively rapid changes in the environment have relied on the ability to modify their behavior during their lifetime. Intelligence, in its simplest form, can be thought of as the flexibility endowed by our genes that allows organisms to adjust their behavior to relatively rapidly changing environments. For some animals, a stable supply of a highly specific food may be predictable (e.g., eucalyptus leaves for the koala or bamboo leaves for the giant panda)—at least until recently. For most animals, however, environments are much less predictable, and their

T.R. Zentall (✉)
Department of Psychology, University of Kentucky
Lexington, KY 40506, USA
e-mail: zentall@uky.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_2,

predisposed eating preferences have had to be much more flexible. For still other animals, the environment is sufficiently unpredictable that it is impossible to for them to be predisposed to know (by genetic means) what food will be available (consider the varied diet of the city-dwelling rat). For these animals to survive, more general (abstract) rules must be available. Rules about what to eat may not be based on the sight or taste of what is ingested but on its consequences. Instead of instructing the animal to eat eucalyptus leaves or to eat a certain class of seeds, these genes instruct the animal that if it feels sick after eating a new food, it should avoid eating more of that food. Such general rules allow for the behavioral flexibility that we call learning.

But there is a price to pay for this added flexibility. The animal must sometimes suffer the consequences of eating something bad. If the novel food is poisonous, the animal may not survive to use its newfound knowledge. The creation and maintenance of a nervous system capable of such learning represents a cost as well. For many animals, the benefits of the capacity for simple associative learning outweigh the cost, but for some animals, the negative consequences of trial and error learning are sufficiently costly that simple learning rules are not enough.

Some animals have found ways to reduce this cost. Rats, which live in highly unpredictable environments, have evolved the ability to learn, in a single experience, the consequences of eating a small amount of a novel food, even when those consequences are experienced hours after the food was ingested (Garcia and Koelling 1966). Rats also have developed the ability to transmit food preferences socially. If a rat experiences the smell of a novel food on the breath of another rat, it will prefer food with that smell over another equally novel food (Galef 1988), and it may also be able to assess the consequences to the other rat of having eaten a novel food (Kuan and Colwill 1997).

But what if this degree of flexibility in learning is still not enough to allow for survival? In the case of humans, for example, our poorly developed sense of smell, our relatively poorly developed gross motor response (e.g., slow running speed), and our relative physical weakness may not have allowed us to hunt competitively with other predators (e.g., large cats). The competition with other animals for food must have come about slowly enough for us to develop weapons and tools, complex forms of communication (language), and complex social structure (allowing for cooperation, teamwork, and reciprocation). According to this view, although our intellect appears to have given us a clear advantage over other animals, its evolution is likely to have emerged because of our relative weakness in other areas. Other animals have compensated for their weaknesses by developing strengths in nonintellectual areas (e.g., the snail compensates for its lack of rapid mobility by building a protective shell around itself). Discussions of animal intelligence often assume, inappropriately, that intelligence is inherently good. In our case, it has turned out to be generally true (at least to the present). For us, intelligence has had a runaway effect on our ability to adapt to change (an effect that Dawkins 1976 calls hypergamy), which has allowed us to produce radical changes in our environment. However, from a biological perspective, in general, intelligence can be viewed as making the best out of a bad situation, or producing a complex solution to problems that other species have often solved in simpler ways. As we evaluate the various intellectual capacities of nonhuman animals, let us try to keep in mind that they have survived quite well (until recently) without the need for our complex intellectual skills.

The Comparative Approach: Two Caveats

First, most people have a vague idea of the relative intelligence of animals. As a general rule, those species that are more like us physically are judged to be more intelligent. But we must be careful in making such judgments because we humans are the ones who are defining intelligent behavior. We make up the rules and the testing procedures, and those tests may be biased in favor of our particular capacities. Isn't it interesting that animals that are more similar to us, that have similar sensory, motor, and motivational systems, just happen to be judged as more intelligent?

Bitterman (1975) has suggested that a relational view of animal learning can be used to correct for peripheral differences in sensory capacity and motor coordination. He suggests that rather than looking for differences in the rate at which different species can learn, we might look at differences, for example, in an animal's ability to learn from the experience of learning. In other words, to what extent can learning facilitate new learning (learning to learn)? Then, using the rate of original learning as a baseline, one can determine the degree to which later learning, presumably involving the same processes, is facilitated. However, this approach is not always possible, and we must be aware that our assessment may be biased by the use of testing procedures not well suited for the species we are studying.

Second, we must guard against the opposite bias—the tendency to interpret behavior as intelligent because of its similarity to intelligent human behavior. In evaluating research addressing the cognitive capacity of animals, we should adopt C. Lloyd Morgan's (1894) position that it is not necessary to interpret behavior as complex (more cognitive) if a simpler (less cognitive) account will suffice. This is the principle of parsimony. Thus, higher-level cognitive interpretations should always be contrasted with simpler, contiguity- and contingency-based, associative-learning accounts. I will start with several classical issues concerned with the nature of learning and intelligence in animals, move to more complex behavior thought to be uniquely human, and end with examples of presumably complex behavior that are likely to be based on simpler predisposed processes.

This review will focus primarily on the flexible behavior of nonprimates because the cognitive behavior of primates is covered elsewhere in this volume, and thus, it will not address several areas of research that have been conducted uniquely with primates, such as analogical reasoning, conservation of volume and mass, language, perspective taking, theory of mind, and deception.

Absolute Versus Relational Learning

One of the most basic cognitive functions involves not being bound to the absolute properties of a stimulus. Although Hull (1943) claimed that learning involves solely the absolute properties of a stimulus, he proposed that animals will appear to respond relationally because they will respond similarly to similar stimuli, a process known as stimulus generalization. Spence (1937) elaborated on this theory by proposing that discrimination learning establishes predictable gradients of excitation (approach) and inhibition (avoidance) that summate algebraically. And this theory of generalization gradient summation can account for a number of phenomena that were formerly explained as relational learning (see Riley 1968). The fact that one sees little discussion of this issue in the modern literature suggests that animals are capable of using either the absolute or relative properties of a stimulus in making discriminations.

Learning to Learn

Can an animal use prior learning to facilitate new learning? That is, can animals learn to learn? If an animal learns a simple discrimination between two stimuli (an S+, to which responses are reinforced, and an S−, to which responses are extinguished) and then, following acquisition, the discrimination is reversed (the S+ becomes S− and the S− becomes S+), and then reversed again, repeatedly, are successive reversals learned faster than earlier reversals? Animals trained on such a serial-reversal task often show improvement within a few reversals, and the rate of improvement can be used as a measure of learning to learn. For example, rats show more improvement than pigeons, and pigeons show more improvement than gold fish (Bitterman and Mackintosh 1969). Mackintosh (1969) attributes these differences in serial-reversal learning to the differential ability of these species to maintain attention to the relevant dimension and ignore irrelevant dimensions.

A different approach to learning to learn is to look for improvement in the rate at which discriminations involving new stimuli are learned. This phenomenon, known as learning set (Harlow 1949), has been studied primarily using visual discriminations with monkeys, but good evidence for such effects has also been found with olfactory discriminations with rats (Slotnick and Katz 1974). In the limit, learning of a new discrimination, or of a reversal, can occur in a single trial. When it does, it is referred to as a win-stay-lose-shift strategy because stimulus choice is completely controlled by the consequences of choice on the preceding trial. One means of developing such a strategy is to learn to forget the consequences of trials prior to the immediately preceding trial. In fact, research has shown that memory for the specific characteristics of the stimuli from prior discriminations does decline as the number of discriminations learned increases (Meyer 1971). Thus, animals approach optimal learning by learning to ignore the effects of all but the most recent experience.

Stimulus Class Formation

Perceptual Classes

Pigeons are remarkably adept at responding selectively to photographs of natural scenes, depending on whether the scene involves a human form (Herrnstein and Loveland 1964) or trees or water (Herrnstein et al. 1976) and those objects need not be anything that they might have actually encountered in their past (e.g., underwater pictures of fish; Herrnstein and deVilliers 1980). To demonstrate that the pigeons do not simply memorize a list of pictures and their appropriate responses, Herrnstein et al. showed that the pigeons would respond appropriately to new examples of the positive and negative stimulus sets.

What is interesting about perceptual classes is that it is difficult to specify what features humans or pigeons use to discriminate members from nonmembers of the perceptual class. However, examination of the kinds of errors made can tell us about the attributes that were used to categorize the exemplars and the similarities in the underlying processes for different species. For example, pigeons make errors similar to those of young children (e.g., they often erroneously assign a picture of a bunch of celery or an ivy-covered wall to the category "tree").

Equivalence Relations

The emergent relations that may arise when arbitrary, initially unrelated stimuli are associated with the same response are often referred to as functional equivalence because they belong to a common stimulus class (see Zentall and Smeets 1996). The best example of equivalence relations in humans is that aspect of language known as semantics—the use of symbols (words) to stand for objects, actions, and attributes. What makes these relations so powerful is what one learns about one member of the stimulus class (i.e., a word) will transfer to others (i.e., the object that it represents). Thus, a child can be told about the varied behavior of dogs (sometimes friendly but not always) without having to actually experience them (and getting bitten). Thus, stimuli that belong to the same stimulus class can be thought of as having the same meaning. The most common procedure for demonstrating the development of functional equivalence in animals involves training on two conditional discriminations. In the first, for example, a red hue (sample) signals that a response to a circle will be reinforced (but not a response to a dot), and a green hue signals that a response to a dot will be reinforced (but not a response to a circle; see Fig. 16.1). In the second conditional discrimination, a vertical line signals that a response to the circle will be reinforced (but not a response to the dot), and a horizontal line signals that a response to the dot will be reinforced (but not a response to the circle). Thus, the red hue and vertical line can be described as meaning *choose the circle* and the green hue and horizontal line as *choose the dot*. This procedure has been referred to as many-to-one matching because training involves the association of two samples with the same comparison stimulus. To show that an emergent relation has developed between the red hue and the vertical

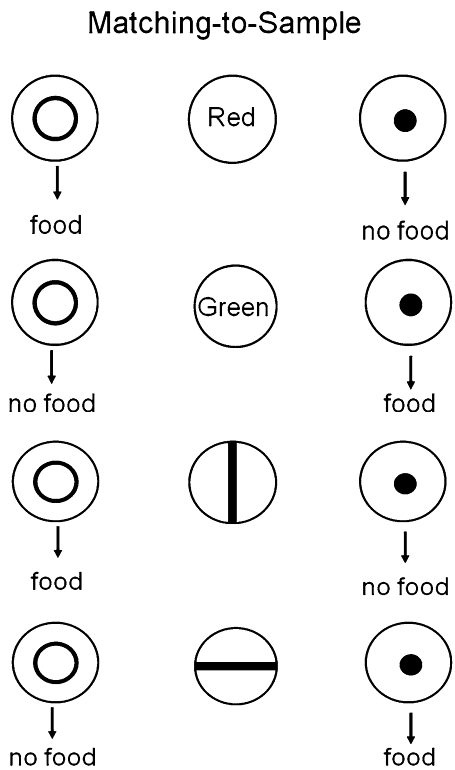


Fig. 16.1 Many-to-one matching training used to show that pigeons will learn that *red* and *vertical* (as well as *green* and *horizontal*) "mean the same thing." If *red* and *green* samples are now associated with new comparison stimuli, *blue* and *white*, respectively, there is evidence that *vertical* and *horizontal lines* are also associated with the *blue* and *white* stimuli, respectively

line and between the green hue and the horizontal line, one can train new associations between one pair of the original samples (e.g., the red and green hues) and a new pair of comparison stimuli (e.g., blue and white hues, respectively). Then on test trials, one can show that emergent relations have developed when, without further training, an animal chooses the blue hue when the sample is a vertical line and chooses the white hue when the sample is a horizontal line (Urcuioli et al. 1989; Wasserman et al. 1992; Zentall 1998).

Although pigeons are not capable of language learning, the ability of small-brained organisms like pigeons to develop arbitrary stimulus classes, the main characteristic of symbolic representation, suggests that this capacity is much more pervasive than once thought.

Memory Strategies

The task most often used to study memory in animals is delayed matching-to-sample, in which following acquisition of matching-to-sample, a delay is inserted between the offset of the sample and the onset of the comparison stimuli (Roberts and Grant 1976). However, the retention functions typically found with this procedure generally greatly underestimate the animal's memory capacity for two reasons. First, in many studies, the novel delay interval is quite similar in appearance to the time between trials. This leads to an ambiguity in the meaning of the delay. When the delay interval and the intertrial interval are made distinctive, the retention functions obtained often provide a very different picture of the animal's memory (Sherburne et al. 1998). Second, the novelty of the delays may result in a generalization decrement that is confounded with memory loss. When pigeons are trained with delays, considerably better memory has been found (Dorrance et al. 2000). Of more interest in the assessment of animal intelligence is an animal's ability to actively affect its memory.

Prospective Processes

Traditionally, animal memory has been viewed as a rather passive process. According to this view, sensory events can leave a trace that may control responding even when the event is no longer present (Roberts and Grant 1976). However, there is evidence that animals can also actively translate or code the representation of a presented stimulus into an expectation of a yet-to-be-presented event (Honig and Thompson 1982). What does it mean to have an expectation of a future event? Imagine a delayed matching task in which vertical- and horizontal-line samples are mapped onto red and green comparison stimuli. During the delay, one can imagine that some representation of the just seen sample stimulus would be remembered. But it is also possible that the sample is translated into a response intention to select one of the comparison stimuli. The ability to use expectations, or prospective coding processes, has important implications for the cognitive capacities of animals. If the expectation of a stimulus, response, or outcome can serve as an effective cue for comparison choice, it suggests that animals may be capable of exerting active control over memory, and in particular, it may suggest they have the capacity for active planning.

The notion of expectancy as an active purposive process can be attributed to Tolman (1932). Although one can say that a dog salivates when it hears a bell because it expects food to be placed in its mouth, the demonstration that an expectation can serve as a discriminative stimulus (i.e., as the basis for making a choice) suggests that the expectancy has additional cognitive properties.

The Differential Outcome Effect If a conditional discrimination or matching task is designed such that a correct response following one sample results in one kind of outcome (e.g., food) and following the other sample results in a different kind of outcome (e.g., water), one can show that acquisition of the conditional discrimination is faster (Trapold 1970) and retention is better when a delay is inserted between the conditional and choice stimuli (Peterson et al. 1980). Furthermore, there is evidence from transfer-of-training experiments that in the absence of other cues, outcome anticipations can serve as sufficient cues for comparison choice. That is, if the original samples are replaced by other stimuli associated with the same differential outcomes, positive transfer has been found (Edwards et al. 1982; Peterson 1984). This line of research indicates that presentation of a sample creates an expectation of a particular kind of outcome and that expectation alone can then serve as the basis for comparison choice. In most cases, the differential outcomes have differential hedonic value (e.g., a high probability of food versus a low probability of food), and it is possible that outcome anticipation can elicit differential emotional states in the animal. But there is also evidence that nondifferentially hedonic events such as the anticipation of a particular neutral stimulus can affect response accuracy (Kelly and Grant 2001; Miller et al. 2009; Williams et al. 1990).

Planning Ahead One of the hallmarks of human cognitive behavior is our ability to consciously plan for the future. Although animals sometimes appear to plan for the future (birds build nests; rats hoard food), these behaviors are likely to be under genetic control. That is, animals do it but it is not likely to be with the expectation of later use. Alternatively, what appears to be future planning just may be the ability to delay reinforcement. To distinguish between planning for the future and learning with a long delay of reinforcement, Suddendorf and Corballis (1997) have suggested that the behavior indicative of planning must occur in the absence of the relevant motivation. Roberts (2002) reported the absence of planning by monkeys. After they had eaten a portion of their daily allotment of food, they threw out of their cage whatever food remained but then requested more food later in the day. However, convincing evidence for planning was reported by Raby et al. (2007). Western scrub jays, which cache food for future use, learned that unpredictably, they would either spend the night in a compartment in which in the morning they would find one kind of food (peanuts) or in a compartment in which they would find a different kind of food (kibble). On test trials, the night before, they were allowed to eat and cache food in either compartment. When they were given peanuts, they tended to cache them in the kibble compartment, and when they were given kibble, they tended to cache them in the peanut compartment (i.e., they cached the food in the compartment in which they would not find that particular food in the morning).

Directed (Intentional) Forgetting

The notion of directed or intentional forgetting is borrowed from human memory research. It implies that memory is an active rather than a passive (automatic) process. Presumably, following presentation, items that human participants are instructed to forget may not be well stored or maintained in memory and, thus, should not be well retained. In a directed forgetting task with animals, for example, pigeons are trained on a matching task, and then a delay of a fixed duration is introduced between the sample and the comparisons. On remember trials, during the delay, the pigeons are cued that there will be a test of sample memory, whereas on forget trials, the pigeons are cued that there will be no test of sample memory. On selected probe trials, the forget cue is presented, but there is a test of sample memory. Matching accuracy on these probe trials is generally below that of remember trials on which there was an expected test of sample memory (Grant 1981). But this design confounds differential motivation on remember and forget trials with sample memory effects because food can be obtained only on remember trials. In a more complex design that controls for motivational effects and that better approximates the human directed forgetting procedure by allowing the animal to reallocate its memory from the sample to an alternative memory on forget trials in training, better evidence for directed forgetting in pigeons has been demonstrated (Roper et al. 1995). Thus, under certain conditions, it appears that animals do have at least some active control over memory processes.

Episodic Memory

Human memory can be identified by the kinds of processes presumed to be involved. Procedural memory involves memory for actions (e.g., riding a bicycle), and it has been assumed that most learned behavior by animals involves this kind of memory. Human declarative memory is assumed to be more cognitive because it involves memory for facts (semantic memory) and memory of personal experiences (episodic memory). Although animals cannot typically describe factual information, their conditional rule-based learning can be thought of as a kind of semantic memory (e.g., if the sample is red, choose the vertical line; if the sample is green, choose the horizontal line). But do animals have episodic memory? Can they retrieve personal experiences or do they simply remember the rules.

Tulving (1972) proposed that an episodic memory should include the what, where, and when of an experience. Clayton and Dickinson (1999) showed that western scrub jays that cached peanuts and wax worms (what) on one side or the other of an ice cube tray (where) learned that their preferred wax worms would be edible after one day but after four days only the peanut would be edible (when; see also Babb and Crystal 2006, for a similar finding with rats). But it can be argued that it is insufficient to retrieve the what, where, and when of an episode because those have been explicitly trained (i.e., they are likely to be semantic or rule-based memories). Instead, better evidence for episodic memory would come from the finding that animals can retrieve information about a past episode when there is no expectation that they will be requested to do so in the future (Zentall et al. 2001). That is, imagine that pigeons are first trained to report the location where they recently pecked (instructions) and then they are trained on an unrelated conditional discrimination in which choice of a vertical line was correct when the sample was blue and choice of the horizontal line was correct when the sample was yellow. Singer and Zentall (2007) found that on probe trials on which following a vertical- or horizontal-line comparison response the pigeons were asked unexpectedly to report the location that they had pecked, they reliably did so. Thus, by either criterion (what-where-when or responding to an unexpected question), pigeons show some evidence of episodic-like memory.

Navigation

Compared to many animals, humans have relatively poor navigational skills. Consider how dependent we are on external supports such as compasses, maps, and more recently global positioning devices. Many animals (e.g., migrating whales, birds, monarch butterflies) can navigate over many hundreds of miles using magnetic fields, chemical gradients, and star patterns. And homing pigeons use a number of these navigational systems including landmarks consisting of natural and man-made geographic features (Lipp et al. 2004).

However, many humans have the ability to imagine a route that they will take and even to imagine how to get to a familiar destination by a novel path. This ability, known as cognitive mapping, consists of mentally knitting together landmarks one has experienced, such that the relation among them can be used to determine a novel path to arrive at a goal. Landmarks are needed to form a cognitive map, but they should not be necessary to use it. Can animals form a cognitive map?

Before trying to answer this question, we need to make an important distinction. Some animals have the remarkable ability to navigate in the absence of landmarks or other external cues. This ability, known as path integration (or dead reckoning), involves the representation of direction and distance one has traveled from a starting point. Dessert ants are particularly adept at path integration as can be shown not only by the direct path that they take to return to their nest after a foraging trip but also by the systematic error incurred if they are displaced just before they attempt to return home (Collette and Graham 2004). The distinction between path integration and cognitive mapping has been a point of controversy. However, under conditions that cannot be accounted for with either landmark use or path integration, there is evidence for the development of a simple cognitive map in rats (Singer et al. 2007) and dogs (Chapuis and Varlet 1987).

Counting

The term numerical competence is often used in animal research because the more common term, counting, carries with it the surplus meaning that accompanies the human verbal labels given to numbers. That this distinction is an arbitrary one, based on limitations of response (output) capacity rather than conceptual ability, is suggested by Pepperberg's (1987) work with generalized verbal number use in an African gray parrot.

An excellent review of the animal counting literature is provided by Davis and Memmott (1982), who conclude that "although counting is obtainable in infra humans, its occurrence requires considerable environmental support" (Davis and Memmott, p. 566). In contrast, Capaldi (1993) concludes that under the right conditions, animals count routinely. In simple but elegant experiments, Capaldi and Miller (1988) demonstrated that following training, rats can anticipate whether they will get fed or not for running down an alley depending solely on the number of successive times they have run down that alley and found food or the absence of food on successive earlier trials.

The difference in the conclusions reached by Davis and Memmott (1982) and by Capaldi and Miller (1988) has general implications for the study of intelligence in animals (including humans). The context in which one looks for a particular capacity may determine whether one will find evidence for it. As noted earlier, because we, as human experimenters, devise the tasks that serve as the basis for the assessment of intelligence, we must be sensitive to the possibility that these tasks may not be optimal for eliciting the behavior we are assessing. That is, much of our view of the evolutionary scale of intelligence may be biased in this way by species differences in sensory, response, and motivational factors.

Reasoning

Reasoning can be thought of as a class of cognitive behavior for which correct responding on test trials requires an inference based on incomplete experience. Although, for obvious reasons, most research on reasoning in animals has been done with higher primates (e.g., chimpanzees), there is evidence that some reasoning-like behavior can be demonstrated in a variety of species.

In its simplest form, the transitive inference task can be described as follows: if A is greater than B (A>B), and B is greater than C (B>C), then it can be inferred that A>C (where the letters A, B, and C represent arbitrary stimuli). Correct responding on this relational learning task requires that an inference be made about the relation between A and C that can only be derived from the two original propositions. To avoid potential problems with end-point effects that could produce a spurious nonrelational solution (i.e., A is always greater, and C is never greater), experimental research typically uses a task that involves four propositions: A > B, B > C, C > D, and D > E, and the test involves the choice between B and D, each of which is sometimes greater and sometimes lesser.

When humans are tested for transitive inference, the use of language allows for the propositions to be completely relational. Relative size may be assigned to individuals identified only by name (e.g., given that Anne is taller than Betty, and Betty is taller than Carol, who is taller, Anne or Carol?). With animals, however, there is no way to present such relational propositions without also presenting the actual stimuli. And if the stimuli differ in observable value (e.g., size), then a correct response can be made without the need to make an inference.

McGonigle and Chalmers (1977) suggested that a nonverbal relational form of the task could be represented by simple simultaneous discriminations in which one stimulus is associated with reinforcement (+) and the other is not (−). A > B can be represented as A + B−, B > C as B + C−, and so on. With four propositions, an animal would be exposed to A + B−, B + C−, C + D−, and D + E−. A is always positive and E is always negative, but B and D, stimuli that were never paired during training, would share similar reinforcement histories. If animals order the stimuli from A is best to E is worst, then B should be preferred over D.

Findings consistent with transitive inference have been reported in research with species as diverse as chimpanzees (Gillan 1981), rats (Davis 1992), and pigeons (Fersen et al. 1991). Although some have argued that these results can be accounted for without postulating that an inference has been made (Couvillon and Bitterman 1992; Fersen et al. 1991; Steirn et al. 1995), transitive inference effects have been found when these presumably simpler mechanisms have been controlled (Lazareva and Wasserman 2006; Weaver et al. 1997). Thus, although it is not clear what mechanism produces it, pigeons clearly show transitive choice that is not produced by differential reinforcement history or differential value that transfers from the positive to the negative stimulus in a simultaneous discrimination.

Taking the Perspective of Others

An organism can take the perspective of another when it demonstrates an understanding of what the other may know. For example, when Susan sees a hidden object moved to a second hidden location after Billy has left the room and Susan understands that Billy will probably look for the object in the first location rather than second, we would say that Susan can take the perspective of Billy or she has a theory of mind because she understands that Billy doesn't know that the object has been moved (see Frye 1993). To demonstrate perspective taking in an animal is a bit more complex because, in the absence of language, theory of mind must be inferred from other behavior (see, e.g., Hare et al. 2001).

Self-recognition

Recognition of the similarity between ourselves and other humans would seem to facilitate perspective taking. If we can recognize ourselves in a mirror, we can see that we are similar to others of our species. Gallup (1970) has shown that not only will chimpanzees exposed to a mirror use it for grooming, but if their face is marked while they are anesthetized, they will use the mirror to explore the mark visually and tactually (i.e., they pass the mark test). Furthermore, both prior experience with the mirror and the presence of the mirror following marking appear to be necessary for mark exploration to occur. Mirror-directed mark exploration appears to occur generally in the great apes (orangutans and perhaps also in gorillas) but not in monkeys even with extensive mirror experience (Gallup and Suarez 1991). However, using the mark test, there is some evidence of self-recognition in dolphins (Reiss and Marino 2001), elephants (Plotnik et al. 2006), and magpies (Prior et al. 2008). Thus, self-recognition appears to occur in several nonprimate species thought to show other kinds of cognitive skills.

Imitation

A more direct form of perspective taking involves the capacity to imitate another (Piaget 1951), especially opaque imitation for which the observer cannot see itself perform the response (e.g., clasping one's hands behind one's back). But evidence for true imitative learning requires that one rule out (or control for) other sources of facilitated learning following observation (see Whiten and Ham 1992; Zentall 1996, 2012). A design that appears to control for artifactual sources of facilitated learning following observation is the two-action procedure based on a method developed by Dawson and Foss (1965). For example, imitation is said to occur if observers, exposed to a demonstrator performing a response in one of two behaviorally different ways, perform the response with the same behavior as their demonstrator. Akins and Zentall (1996) trained Japanese quail demonstrators to either step on a treadle or peck the treadle for food reinforcement. When observer quail were exposed to one or the other demonstrator, they matched the behavior of their demonstrator with a high probability (see also Zentall et al. 1996, for similar evidence with pigeons). Furthermore, there is some evidence that pigeons can imitate a sequence of two responses, operating a treadle (by stepping or pecking) and pushing a screen (to the left or to the right; Nguyen et al. 2005), an example of what Byrne and Russon (1998) refer to as program-level imitation.

If Piaget is correct, the ability to imitate requires the ability to take the perspective of another. But children do not develop the ability to take the perspective of another until they are 5–7 years old, yet they are able to imitate others at a much earlier age. Furthermore, if pigeons and Japanese quail can imitate, it is unlikely that they do so by taking the perspective of the demonstrator, in the sense that Piaget implied. Thus, although cognitively interesting, imitation may not provide evidence for the kind of cognitive behavior implied by perspective taking.

What Animals Can Tell Us About Human Reasoning

I have saved for last the discussion of several lines of research with animals directed to biases and heuristics characteristic of humans that appear to be somewhat irrational or at least suboptimal. The results of these studies are important, not so much because of their implications for animals, but primarily for their implications for how we interpret human behavior. That is, if other animals have these same biases, then the basis for those biases does not depend on language or human culture as is sometimes proposed.

Cognitive Dissonance

One of these biases has to do with a phenomenon extensively studied in humans called cognitive dissonance. Cognitive dissonance is the discomfort that comes when there is a discrepancy between one's beliefs and one's behavior. For example, if one believes that one should tell the truth, one is likely to feel dissonance on occasions when one fails to do so. That dissonance may be resolved by deciding that there are some conditions under which lying is appropriate or the person lied to may have deserved it. Cognitive dissonance presumably comes about because of a need to be consistent or to avoid being labeled a hypocrite. Does this represent a kind of social intelligence? And if so, would nonhuman animals show a similar effect? But how would one go about asking this question of animals?

One approach involves a version of cognitive dissonance called justification of effort (Aronson and Mills 1959). In their study, undergraduates, who underwent an unpleasant initiation to become part of a group, reported that they wanted to join the group more than those who underwent a less unpleasant initiation. It is assumed that those individuals who underwent an unpleasant initiation gave more value to membership in the group to justify undergoing the unpleasantness.

The justification of effort design allows for a direct test of cognitive dissonance in animals. For example, if on some trials a pigeon has to work hard to receive signal A that says food is coming and on other trials the pigeon does not have to work hard to receive signal B that says the same food is coming, will the pigeon show a preference for signal A over signal B? Several studies have shown that they will (e.g., Clement et al. 2000; Kacelnik and Marsh 2002). But is this cognitive dissonance? Do animals need to justify to themselves why they worked harder for one signal than the other?

Alternatively, we have suggested that this choice behavior results from the contrast between the relatively negative emotional state of the organism at the end of the effort and upon presentation of the signal (Zentall and Singer 2007). That difference would be greater when more effort was involved. Thus, the subjective value of the signal for reinforcement might be judged to be greater. Contrast provides a more parsimonious account of the pigeons' choice behavior. Could contrast also be involved in the similar behavior shown by humans? This possibility should be examined by social psychologists.

Maladaptive Gambling Behavior

Humans often gamble (e.g., play the lottery) even though the odds against winning are very high. This behavior may be attributable to an inaccurate assessment of the probability of winning, perhaps resulting in part from public announcements of the winners but not the losers (an availability heuristic). Would animals show a similar kind of maladaptive gambling behavior? According to optimal foraging theory, they should not because such inappropriate behavior should have been selected against by evolution. Furthermore, if the choice is to have any meaning for the animal, it would have to have experienced the probability associated with winning (reinforcement) and that should reduce the likelihood that the animal would not be able to assess the probability of winning and losing. However, we have recently found conditions under which pigeons will prefer an average of 2 pellets of food over a predictable 3 pellets of food (Zentall and Stagner 2011). The procedure is as follows: If the pigeon chooses the left alternative, on 20 % of the trials, a green stimulus appears and is followed by 10 pellets of food. The remainder of the time it chooses the left alternative; a red stimulus appears and is never followed by food. Thus, on average the pigeon receives 2 pellets of food for choosing that alternative. If the pigeon chooses the right alternative, it received either a blue or a yellow stimulus but in either can it receives 3 pellets of food. Curiously, the pigeons prefer the left alternative overwhelmingly over the right alternative, and they do so in spite of the fact that they would get 50 % more food for choosing the right alternative.

This result suggests that gambling behavior is likely to have a simple biological basis, and although social and cognitive factors may contribute to human gambling behavior, the underlying mechanism is likely to be simpler. The mechanisms responsible for this suboptimal behavior appear to involve the enhanced effect of the signal for the large magnitude of reinforcer (the 10-pellet "jackpot") and the reduced effect of the signal for nonreinforcement with training (Stagner et al. 2012). This account appears to be consistent with research with humans that has found that gamblers overvalue wins in spite of their low probability of occurrence and they give too little negative value to their losses in spite of their high probability of occurrence (Blanco et al. 2000).

Sunk Cost

The sunk cost effect occurs when one allows an amount of money, time, or other resource already invested to affect one's decision to invest more resources. For example, one may sit through a film that one does not like because to leave would be to waste one's investment of the price of the ticket. But in so doing, one is spending additional resources, one's time, and there is no way to recoup the money already invested. Similarly, one may choose to continue with a failing business because of the past investment one has made in it. This phenomenon comes under the general rubric of prospect theory (Kahneman and Twersky 1979) which suggests that humans will take greater risks to avoid a loss than to obtain a gain. The question is to what extent does this behavior stem from the cultural tenet to avoid wasting resources and to complete what one starts. If one could show that animals show the same behavior, it would suggest that sunk cost is a general phenomenon that has basic behavioral origins.

In fact, evidence for sunk cost has been found in pigeons (Navarro and Fantino 2005; Pattison et al. 2012). For example, pigeons learn that pecking a green light requires 30 pecks, whereas pecking a red light requires only 10 pecks. They then learn that after pecking the green light a number of times (that varied from trial to trial), they would be able to choose to continue with the green light (to complete the 30 pecks) or switch to the red light for which 10 pecks were required. Results indicated that the pigeons often choose to return to the green light even when 20 more pecks are required. Thus, pigeons show a sunk cost effect that is very similar to that shown by humans. Why pigeons show the sunk cost effect is not clear. One can speculate that it arises from the fact that in nature switching to a different patch often involves uncertainty, some travel time, and possible danger, but one can certainly conclude that culture and language are not necessary components.

When Less Is More

When humans are asked to judge the value of a set of objects of excellent quality, they often give it higher value than those same objects with the addition of some objects of lesser quality (Hsee 1998). This bias is an example of the affect heuristic in which it appears that the average quality of a set is used to determine the value of the set rather than the quantity of items in the set. The phenomenon has become known as a less is more effect. It has been found when humans are asked to judge the value of sets of dishes and sets of baseball cards (Hsee 1998), and it also has been found when academics are asked to judge the quality of a curriculum vita (Hayes 1983). For example, a vita with three publications in excellent journals is judged better than one with the same three publications in excellent journals plus six more in lesser quality journals.

Recent evidence suggests that even pigeons are susceptible to this bias. We found that pigeons will work for dried peas and dried milo seeds, but when given a choice between the two, they prefer the peas. However, when they are given a choice between a pea and a pea together with a milo seed, they prefer the pea alone (Zentall et al. 2013). Apparently, the pigeons too are averaging the high-quality pea with the lower-quality milo seed and value the pair less than the pea by itself (see also Kralik et al. 2012). The basis of this bias may originate in the need to make rapid decisions, presumably because of intense competition from conspecifics and the possibility of predation, and they use it in the laboratory even when speed is not a factor. Once again, the fact that other animals show this suboptimal choice indicates that the bias is probably not dependent on human cultural influence.

Conclusions

The broad range of positive research findings that have come from investigating the cognitive abilities of animals suggests that many of the “special capacities” attributed to humans may be more quantitative than qualitative. In the case of many cognitive learning tasks, once we learn how to ask the question appropriately (i.e., in a way that is accommodating to the animal), we may often be surprised with the capacity of animals to use complex relations.

In evaluating the animal (and human) intelligence literature, we should be sensitive to both overestimation of capacity (what appears to be higher-level functioning in animals that can be accounted for more parsimoniously at a lower level; see Zentall 1993) and underestimation of capacity (our bias to present animals with tasks convenient to our human sensory, response, and motivational systems). Underestimation can also come from the difficulty in providing animals with task instructions as one can quite easily do with humans (see Zentall 1997). The accurate assessment of animal intelligence will require vigilance, on the one hand, to evaluate cognitive functioning against simpler accounts and, on the other hand, to determine the conditions that will maximally elicit the animal's cognitive capacity.

The study of human biases by examining animals for the presence of similar phenomena in animals can also help us to determine that simpler mechanisms are involved. Thus, the study of animal intelligence can inform us not only of the cognitive abilities of animals but also can suggest the bases of certain human phenomena thought to have complex social origins.

Acknowledgments Preparation of this chapter was made possible by grant MH 63726 from the National Institute of Mental Health and grant HD60996 from the National Institute of Child Health and Human Development.

References

Akins, C., & Zentall, T. R. (1996). Evidence for true imitative learning in Japanese quail. *Journal of Comparative Psychology, 110*, 316–320.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology, 59*, 177–181.

Babb, S. J., & Crystal, J. D. (2006). Discrimination of what, when, and where is not based on time of day. *Learning & Behavior, 34*, 124–130.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science, 188*, 699–709.

Bitterman, M. E., & Mackintosh, N. J. (1969). Habit reversal and probability learning: Rats, birds, and fish. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 163–185). New York: Academic.

Blanco, C., Ibáñez, Sáiz-Ruiz, A., Blanco-Jerez, J., & Nunes, E. V. (2000). Epistemology, pathophysiology, and treatment of pathological gambling. *CNS Drugs, 13*, 397–407.

Bryne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences, 21*, 667–721.

Capaldi, E. J. (1993). Animal number abilities: Implications for a hierarchical approach to instrumental learning. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence* (pp. 191–209). Hillsdale: Erlbaum.

Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 3–17.

Chapuis, N., & Varlet, C. (1987). Short cuts by dogs in natural surroundings. *Quarterly Journal of Experimental Psychology, 39*, 49–64.

Clayton, N. S., & Dickinson, A. (1999). Scrub jays (*Aphelocoma coerulescens*) remember the relative time of caching as well as the location and content of their caches. *Journal of Comparative Psychology, 113*, 403–416.

Clement, T. S., Feltus, J., Kaiser, D. H., & Zentall, T. R. (2000). "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review, 7*, 100–106.

Collette, T. S., & Graham, P. (2004). Animal navigation: Path integration, visual landmarks and cognitive maps. *Current Biology, 14*, 475–477.

Couvillon, P. A., & Bitterman, M. E. (1992). A conventional conditioning analysis of "transitive inference" in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 308–310.

Davis, H. (1992). Transitive inference in rats (*Rattus norvegicus*). *Journal of Comparative Psychology, 106*, 342–349.

Davis, H., & Memmott, J. (1982). Counting behavior in animals: A critical evaluation. *Psychological Bulletin, 92*, 547–571.

Dawkins, R. (1976). *The selfish gene*. New York: Oxford University Press.

Dawson, B. V., & Foss, B. M. (1965). Observational learning in budgerigars. *Animal Behaviour, 13*, 470–474.

Dorrance, B. R., Kaiser, D. H., & Zentall, T. R. (2000). Event duration discrimination by pigeons: The choose-short effect may result from retention-test novelty. *Animal Learning & Behavior, 28*, 344–353.

Edwards, C. A., Jagielo, J. A., Zentall, T. R., & Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching-to-sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes, 8*, 244–259.

Fersen, L. V., Wynne, C. D. L., Delius, J. D., & Staddon, J. E. R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 334–341.

Frye, D. (1993). Causes and precursors of children's theory of mind. In D. F. Hay & A. Angold (Eds.), *Precursors and causes of development and psychopathology*. Chichester: Wiley.

Galef, B. G., Jr. (1988). Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. In T. R. Zentall & B. G. Galef Jr. (Eds.), *Social learning: Psychological and biological perspectives* (pp. 3–28). Hillsdale: Erlbaum.

Gallup, G. G. (1970). Chimpanzees self-recognition. *Science, 167*, 86–87.

Gallup, G. G., & Suarez, S. D. (1991). Social responding to mirrors in rhesus monkeys: Effects of temporary

mirror removal. *Journal of Comparative Psychology, 105*, 376–379.
Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science, 4*, 123–124.
Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes, 7*, 150–164.
Grant, D. S. (1981). Stimulus control of information processing in pigeon short-term memory. *Learning and Motivation, 12*, 19–39.
Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour, 61*, 139–151.
Harlow, H. F. (1949). The formation of learning sets. *Psychological Review, 56*, 51–65.
Hayes, S. C. (1983). When more in less: Quantity versus quality of publications in the evaluation of vitae. *The American Psychologist, 38*, 1398–1400.
Herrnstein, R. J., & deVilliers, P. A. (1980). Fish as a natural category for people and pigeons. *Psychology of Learning and Motivation, 14*, 59–95.
Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science, 146*, 549–551.
Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2*, 285–301.
Honig, W. K., & Thompson, R. K. R. (1982). Retrospective and prospective processing in animal working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 239–283). Orlando: Academic.
Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making, 11*, 107–121.
Hull, C. L. (1943). *Principles of behavior*. New York: Appleton.
Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour, 63*, 245–250.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
Kelly, R., & Grant, D. S. (2001). A differential outcomes effect using biologically neutral outcomes in delayed matching-to-sample with pigeons. *Quarterly Journal of Experimental Psychology, 54B*, 69–79.
Kralik, J. D., Xu, E. R., Knight, E. J., Khan, S. A., & Levine, J. W. (2012). When less is more: Evolutionary origins of the affect heuristic. *PLoS One, 7*(10), e46240. doi:10.1371/journal.pone.0046240.
Kuan, L. A., & Colwill, R. (1997). Demonstration of a socially transmitted taste aversion in the rat. *Psychonomic Bulletin & Review, 4*, 374–377.
Lazareva, O. F., & Wasserman, E. A. (2006). Effect of stimulus orderability and reinforcement history on transitive responding in pigeons. *Behavioural Processes, 72*, 161–172.
Lipp, H.-P., Vyssotski, A. L., Wolfer, D. P., Renaudineau, S., Savini, M., Tröster, G., & Dell'Omo, G. (2004). Pigeon homing along highways and exits. *Current Biology, 14*, 1239–1249.
Mackintosh, N. J. (1969). Comparative studies of reversal and probability learning: Rats, birds, and fish. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 137–162). New York: Academic.
McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? *Nature, 267*, 694–696.
Meyer, D. R. (1971). Habits and concepts of monkeys. In L. E. Jarrard (Ed.), *Cognitive processes of nonhuman primates* (pp. 83–102). New York: Academic.
Miller, H. C., Friedrich, A. M., Narkavic, R. J., & Zentall, T. R. (2009). A differential outcomes effect using hedonically-nondifferential outcomes with delayed matching-to-sample by pigeons. *Learning & Behavior, 37*, 161–166.
Morgan, C. L. (1894). *An introduction to comparative psychology*. London: Scott.
Navarro, A. D., & Fantino, E. (2005). The sunk cost effect in pigeons and humans. *Journal of the Experimental Analysis of Behavior, 83*, 1–13.
Nguyen, N. H., Klein, E. D., & Zentall, T. R. (2005). Imitation of two-action sequences by pigeons. *Psychonomic Bulletin & Review, 12*, 514–518.
Pattison, K. F., Zentall, T. R., & Watanabe, S. (2012). Sunk cost: Pigeons (*Columba livia*) too show bias to complete a task rather than shift to another. *Journal of Comparative Psychology, 126*, 1–9.
Pepperberg, I. M. (1987). Interspecies communication: A tool for assessing conceptual abilities in an African Grey parrot. In G. Greenberg & E. Tobach (Eds.), *Language, cognition, and consciousness: Integrative levels* (pp. 31–56). Hillsdale: Erlbaum.
Peterson, G. B. (1984). How expectancies guide behavior. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 135–148). Hillsdale: Erlbaum.
Peterson, G. B., Wheeler, R. L., & Trapold, M. A. (1980). Enhancement of pigeons' conditional discrimination performance by expectancies of reinforcement and nonreinforcement. *Animal Learning & Behavior, 8*, 22–30.
Piaget, J. (1951). *Play, dreams, and imitation in childhood*. New York: W. W. Norton.
Plotnik, J. M., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences, 103*, 17053–17057.
Prior, H., Schwarz, A., & Gunturkun, O. (2008). Mirror-induced behavior in the magpie (*Pica pica*): Evidence of self-recognition. *PLoS Biology, 6*(8), e202. doi:10.1371/journal.pbio.0060202.
Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Empirical evaluation of mental time travel. *Behavioral Brain Sciences, 30*, 330–331.
Reiss, D., & Marino, L. (2001). Self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences, 98*, 5937–5942.
Riley, D. A. (1968). *Discrimination learning*. Boston: Allyn & Bacon.

Roberts, W. A. (2002). Are animals stuck in time? *Psychological Bulletin, 128*, 473–489.

Roberts, W. A., & Grant, D. S. (1976). Studies of short-term memory in the pigeon using the delayed matching-to-sample procedure. In D. L. Medin, W. A. Roberts, & R. T. Davis (Eds.), *Processes of animal memory* (pp. 79–112). Hillsdale: Erlbaum.

Roper, K. L., Kaiser, D. H., & Zentall, T. R. (1995). Directed forgetting in pigeons: The role of alternative memories in the effectiveness of forget cues. *Animal Learning & Behavior, 23*, 280–285.

Sherburne, L. M., Zentall, T. R., & Kaiser, D. H. (1998). Timing in pigeons: The choose-short effect may result from "confusion" between delay and intertrial intervals. *Psychonomic Bulletin & Review, 5*, 516–522.

Singer, R. A., Abroms, B. D., & Zentall, T. R. (2007). Formation of a simple cognitive map by rats. *International Journal of Comparative Psychology, 19*, 417–425.

Singer, R. A., & Zentall, T. R. (2007). Pigeons learn to answer the question 'Where did you just peck?' and can report peck location when unexpectedly asked. *Learning & Behavior, 35*, 184–189.

Slotnick, B. M., & Katz, H. M. (1974). Olfactory learning-set formation in rats. *Science, 185*, 796–798.

Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review, 44*, 430–444.

Stagner, J. P., Laude, J. R., & Zentall, T. R. (2012). Pigeons prefer discriminative stimuli independently of the overall probability of reinforcement and of the number of presentations of the conditioned reinforcer. *Journal of Experimental Psychology: Animal Behavior Processes, 38*, 446–452.

Steirn, J. N., Weaver, J. E., & Zentall, T. R. (1995). Transitive inference in pigeons: Simplified procedures and a test of value transfer theory. *Animal Learning & Behavior, 23*, 76–82.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs, 123*, 133–167.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.

Trapold, M. A. (1970). Are expectancies based on different reinforcing events discriminably different? *Learning and Motivation, 1*, 129–140.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–403). New York: Academic.

Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., & Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes, 15*, 264–273.

Wasserman, E. A., DeVolder, C. L., & Coppage, D. J. (1992). Non-similarity based conceptualization in pigeons via secondary or mediated generalization. *Psychological Science, 6*, 374–379.

Weaver, J. E., Steirn, J. N., & Zentall, T. R. (1997). Transitive inference in pigeons: Control for differential value transfer. *Psychonomic Bulletin and Review, 4*, 113–117.

Whiten, A., & Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. *Advances in the Study of Behavior, 21*, 239–283.

Williams, D. A., Butler, M. M., & Overmier, J. B. (1990). Expectancies of reinforcer location and quality as cues for a conditional discrimination in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 16*, 3–13.

Zentall, T. R. (1993). Animal cognition: An approach to the study of animal behavior. In T. R. Zentall (Ed.), *Animal cognition: A tribute to Donald A. Riley* (pp. 3–15). Hillsdale: Erlbaum.

Zentall, T. R. (1996). An analysis of imitative learning in animals. In C. M. Heyes & B. G. Galef Jr. (Eds.), *Social learning and tradition in animals* (pp. 221–243). New York: Academic.

Zentall, T. R. (1997). Animal memory: The role of instructions. *Learning and Motivation, 28*, 248–267.

Zentall, T. R. (1998). Symbolic representation in pigeons: Emergent stimulus relations in conditional discrimination learning. *Animal Learning & Behavior, 26*, 363–377.

Zentall, T. R. (2012). Perspectives on social learning. *Journal of Comparative Psychology, 126*, 114–128. doi:10.1037/a0025381.

Zentall, T. R., & Singer, R. A. (2007). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than less aversive event. *Journal of the Experimental Analysis of Behavior, 88*, 131–149.

Zentall, T. R., & Smeets, P. M. (Eds.). (1996). *Stimulus class formation in humans and animals*. Amsterdam: North Holland.

Zentall, T. R., & Stagner, J. P. (2011). Maladaptive choice behavior by pigeons: An animal analog of gambling (sub-optimal human decision making behavior). *Proceedings of the Royal Society B: Biological Sciences, 278*, 1203–1208.

Zentall, T. R., Sutton, J. E., & Sherburne, L. M. (1996). True imitative learning in pigeons. *Psychological Science, 7*, 343–346.

Zentall, T. R., Clement, T. S., Bhatt, R. S., & Allen, J. (2001). Episodic-like memory in pigeons. *Psychonomic Bulletin & Review, 8*, 685–690.

Zentall, T. R., Laude, J. R., Daniels, C. W., & Case, J. P. (2013, August). When less means more. Paper presented at the meeting of the American Psychological Association, Honolulu.

3 Intelligence in Nonhuman Primates

Tara Mandalaywala, Christine Fleener, and Dario Maestripieri

Introduction

Intelligence is a multidimensional construct that can include a wide range of cognitive abilities. Historically, intelligence in humans and other animals has been measured and defined by flexibility in problem solving, learning, memory, reasoning, abstract thinking, planning, and communication and language. Humans are highly proficient in all of these domains, but among nonhuman animals there is more variability, and some species perform well on one or a few measures of intelligence but poorly on others.

With such variability in animal intelligence, we must ask why and how intelligence evolved and why species differences in intelligence exist. Intelligence is a metabolically costly brain function, as it is associated with growing larger brains over longer developmental periods. Most life forms on Earth show highly sophisticated adaptations to their environment but little or no general intelligence, suggesting that such general intelligence is not a prerequisite for survival or successful reproduction on our planet. However, intelligence abilities have emerged in many different evolutionary lineages, indicating that the benefits of intelligence can outweigh its costs under certain selective pressures. In the primate lineage in particular, general intelligence has evolved to a greater degree than in almost any other taxonomic group, suggesting that any effort to understand the evolution of human intelligence must take into account broad evolutionary trends within the Primate order as well as similarities and differences between humans and nonhuman primates (NHPs).

The order Primates is 60–85 million years old and is split into two main groups: prosimians and anthropoids. Prosimians, which include lemurs, lorises, and tarsiers, are the most ancestral primates. Anthropoids include monkeys, apes, and hominins and are split into platyrrhines, flat-nosed New World monkeys of South and Central America, and catarrhines, narrow-nosed Old World monkeys and apes of Africa and southeast Asia. Many prosimians and New World monkeys

T. Mandalaywala, Ph.D.(✉)
Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

Department of Comparative Human Development, Institute for Mind and Biology, University of Chicago, 940 E. 57th Street, Chicago, IL 60637, USA
e-mail: tmandalaywala@uchicago.edu

C. Fleener, M.A.
Department of Comparative Human Development, Institute for Mind and Biology, University of Chicago, 940 E. 57th Street, Chicago, IL 60637, USA
e-mail: cfleener@uchicago.edu

D. Maestripieri, Ph.D.
Comparative Human Development, Evolutionary Biology, and Neurobiology, The University of Chicago, 5730 S. Woodlawn Avenue, Chicago, IL 60637, USA
e-mail: Dario@uchicago.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_3,

communicate mainly through olfaction or simple vocalizations, while Old World monkeys and apes rely more heavily on complex visual, vocal, and tactile signals than on olfaction. These forms of communication in Old World monkeys and apes demand more behavioral flexibility and as such involve more complex cognitive abilities. Using comparative studies, we can see how intelligence has evolved in the order Primates, with prosimians showing the least cognitive complexity, New World monkeys showing greater complexity, Old World monkeys showing even greater complexity than New World monkeys, and a significant jump in the cognitive abilities of apes and humans (Deaner et al. 2006).

Despite their cognitive differences, all primates are characterized by having large brains and, in particular, a large neocortex, which is involved in sensory perception, motor commands, spatial reasoning, and conscious thought. Primates develop more slowly and live longer than most mammals, which facilitates a long process of brain growth and maintenance of brain plasticity. Large, plastic brains allow for advanced, flexible cognitive abilities. Using their intelligence, primates can keep track of elaborate dominance hierarchies and recognize conspecifics and kin. They can manipulate, deceive, or cooperate. Some can solve problems in novel ways, make tools, and even be taught to use symbols and language.

To explore how large brains and intelligence arose in primates, it is useful to consider the broad range of ecological and social pressures primates experience in their environment. Some primate species are mostly arboreal and their diet consists mainly of leaves, fruit, or insects; others are mostly terrestrial and their diet can include grass, leaves, fruit, or meat. Some primates live in social groups made up of hundreds of individuals, while others live more solitary lifestyles. As we will see in the next section, some theories concerning the evolution of primate intelligence emphasize the importance of ecological pressures, while others focus on the role of complex social environments.

The Evolution of Intelligence in Primates

The two predominant theories advanced to explain the evolution of intelligence in primates emphasize the cognitive challenges imposed by (1) feeding ecology or (2) the social environment. These are known as the ecological intelligence and social intelligence hypotheses, respectively. Although recent research more strongly supports the social intelligence hypothesis, it is possible that both theories, and possibly other not yet developed theories, are partially correct (e.g., Reader et al. 2011; Holekamp 2007). Both the ecological and social intelligence hypotheses share the assumption that brain tissue is one of the most metabolically expensive tissues both to grow and to maintain (e.g., Aiello and Wheeler 1995). In human infants, brain tissue accounts for approximately 60 % of the basal metabolic rate, and even in adulthood brain tissue requires a large proportion of basal metabolism, around 20 % (Martin 2013).

Ecological Intelligence Hypothesis

The ecological intelligence hypothesis (Milton 1981) proposes that foraging challenges posed by the environment were the impetus for complex cognitive evolution in primates. Most primates consume a predominantly plant-based diet, with different species of primates consuming different types of plants (e.g., folivorous primates consume mostly leaves, whereas frugivorous primates consume a mixture of leaves and fruits). Different plant species have different distributions in space and time, and each type of plant imposes different foraging demands on the primate species that consumes them. For example, leaves tend to be widely distributed but are of low nutritional quality; thus primates must consume more leaves to gain the same energy intake obtained from consuming fewer plants of higher quality, such as fruit (see Fish and Lookwood 2003 for evidence of dietary constraints on brain size). Fruits tend to be patchily distributed in both space and time, imposing a unique challenge on the frugivorous species that depend on them. This environmental complexity requires that frugivorous species possess a suite of strong cognitive skills to forage efficiently, and the ecological intelligence hypothesis proposes that species facing the patchiest and most challenging foraging environments should be the most cognitively advanced (Milton 1981). Consistent with this hypothesis, frugivores have a larger cranial capacity than folivores (Clutton-Brock and Harvey 1980).

In an empirical test of the ecological intelligence hypothesis, Platt et al. (1996) compared Wied's marmosets and golden lion tamarins on spatial and visual memory tasks. Although the diets of these two species overlap, significant differences in spatial and temporal qualities of their foraging led Platt et al. (1996) to hypothesize that the tamarins would outperform the marmosets on the cognitive tasks. In accordance with the ecological intelligence hypothesis, Platt et al. (1996) found that, after a long delay (24 and 48 h), the tamarins outperformed the marmosets, perhaps because of the increased memory demands on the tamarins when seeking out food sources that had more complex spatial and temporal distribution than the food sources of the marmosets.

Parker and Gibson (1979) proposed that it was not the complexity of the food environment but rather the complexity of the foraging tasks posed by the foods consumed that selected for increased intelligence in primates. Those species that engage in "extractive foraging," in which a food object is not immediately or obviously perceptible and must be extracted from a nonfood item, would possess enhanced intelligence relative to those species that only consume readily perceptible foods. Parker and Gibson (1979) suggest that these complex foraging tasks present two cognitively taxing problems: the ability to understand that a nonfood item might contain an edible substance and the ability to engage in complex object manipulation. Evidence in favor of the extractive foraging hypothesis came in the form of higher rates of tool use and object manipulation of food items in species that displayed more cognitively complex behaviors, such as chimpanzees and capuchins (Parker and Gibson 1977; reviewed in King 1986). However, both the ecological intelligence and the extractive foraging hypotheses are problematic as they fail to explain cognitive behavior in a number of species (see King 1986 for a review of species that are considered to have high intelligence but do not engage in extractive foraging).

Social Intelligence Hypothesis

The social intelligence hypothesis postulates that increased group size and the demands imposed by increased social complexity selected for bigger brains and enhanced cognition (Dunbar 1998; Humphrey 1976). Although group living confers many advantages (e.g., Silk 2007), it also has drawbacks such as increased competition for resources, which includes both mates and food. Both Humphrey (1976) and Dunbar (1998) proposed that primates solve these social challenges with the use of flexible cognitive strategies in real time, rather than evolved "rules-of-thumb" and heuristics.

The social intelligence hypothesis predicts that all social species should show enlarged brain size relative to more solitary species because of the increased demands of navigating a more complex social environment and keeping track of an exponentially greater number of relationships. In line with this prediction, Dunbar (1998) found that neocortex size correlates with social group size but does not correlate with any measure proposed by the ecological intelligence or extractive foraging hypotheses. The neocortex has increased in size exponentially across primate evolution (Passingham 1982), and as brain tissue is prohibitively expensive to grow and maintain, the selection pressures for a dramatic increase in neocortex size must have outweighed the metabolic cost (Dunbar 1998) suggesting the particular importance of the neocortex in primates. Moreover, neocortex size correlates with a variety of measures of social complexity in primates such as the number of females in a group, the size of grooming groups, as well as the amount of tactical deception and social learning (reviewed in Dunbar and Shultz 2007; Reader and Laland 2002; but see Grueter et al. 2013). For example, research in various species of lemurs has found that species that live in larger groups are better at tasks of transitive inference than species that live in smaller groups (MacLean et al. 2008).

A related theory, the Machiavellian intelligence hypothesis, suggests that it was the need to engage in social manipulation that helped drive increased cognitive ability (reviewed in Whiten and Byrne 1997). Evidence for this hypothesis comes from studies of deception. Tactical deception "occurs when an individual is able to use an 'honest' act from his normal repertoire in a different context to mislead familiar individuals" (Whiten and Byrne 1988), and it is not found in all primate species; it is distributed unevenly with greater rates of deception occurring in the great apes, and virtually no instances of deception found in prosimians thus far (Byrne and Whiten 1992). Moreover, there is a relationship between neocortical volume and frequency of deception in primates (Byrne and Corp 2004). Implicit in the definition of deception is the application of a behavior to novel situations, emphasizing the utility of flexibility in thinking (e.g., applying a behavior or response to a novel situation when appropriate), not just increased memory or visuospatial skills. Flexible thinking is often seen as a hallmark of true intelligence. Unfortunately, flexible thinking is not well studied. Some hypothesize that one should be able to distinguish hard-wired intelligence—that is, intelligence arising as a result of specific challenges in the environment—from higher-order intelligence and flexible thinking, in which a subject can solve a completely novel problem (see also cultural intelligence: van Schaik et al. 2012).

Just as there are problems with the ecological intelligence hypothesis, some question the necessity of enhanced cognition to deal with social complexity (e.g., Barrett et al. 2007). Having reviewed the main theories as to how and why enhanced intelligence may have evolved in the order Primates, we will now review the many different ways in which nonhuman primates display their intelligence across physical and social domains, in the field and in the laboratory.

Physical Cognition

Primates navigate many different types of environment, and as such, have evolved flexibility in how they process and interact with their physical world. Through both natural selection and experience, primates have acquired an understanding of how objects should behave in space and time and are sometimes able to apply simple principles of physics and mathematics to solving specific problems. Physical cognition refers to the expectation that objects in an environment will follow consistent laws in their behavior and interaction. For example, without knowing anything about an object, we expect that all parts of that object can only trace one connected path through space and time, that two distinct objects cannot occupy the same space at the same time, that the object will not deform as it moves, and that two distinct objects will move together if and only if they make contact in some way. These physical concepts are termed continuity, boundedness, rigidity, and no action at a distance, respectively (Spelke 1990), and there is evidence to show that NHPs comprehend these concepts as well (see below).

Both apes and monkeys have shown the ability to represent the existence and movements of objects even when they cannot be seen (capuchins, chimpanzees: Mireille et al. 1976; Rhesus macaques: de Bois and Novak 1994; orangutans, squirrel monkeys: de Bois et al. 1998), but this research reveals qualitative differences in how apes and monkeys understand physical concepts, illuminating how physical cognition has evolved in primates. In visible displacement tasks where an object moves behind an occluder, both apes and monkeys are able to track the object when it cannot be seen, but in invisible displacement tasks, where the object is first placed in a container and then moved behind an occluder, only apes can track the object. Other research shows that NHPs can understand that a solid object cannot move through another solid object (rhesus macaques: Santos and Hauser 2002), but this research again reveals qualitative differences in how the concept of solidity is understood. For example, while monkeys are surprised when an object appears to fall through a solid barrier, they are not surprised if the object appears to roll down a ramp and through a solid barrier. This error may be due to a "gravity bias" that is also present in infant humans, where most attention is given to an item that is dropped with the expectation that it will fall on the surface directly below where it was dropped (Hood et al. 1999). When multiple forces act on a falling object (i.e., ramps, tubes), both adult NHPs and infant humans are unable to track the object's movement in space and time (Santos 2004). This comparison of infant humans and nonhuman primates suggests similar foundational cognitive circuitry.

Other research shows that NHPs are capable of recognizing that objects cannot move without first being contacted by another object (cotton-top tamarins: Hauser 1998), can differentiate between adequate and inadequate physical support (chimpanzees: Cacchione and Horst 2004), and can infer the location of rewards based on the effect of weight (chimpanzees: Hanus and Call 2008). Having a foundation of these simple physical concepts allows primates to more quickly negotiate their environment and inform their problem solving. While there is extensive evidence showing that NHPs comprehend many physical properties, the question is *how* they are capable of applying this information to the way they reason about their physical world.

Tool Use and Causal Understanding

The primary morphological feature that distinguishes primates from most other animals is their hands. Primates evolved manual dexterity for arboreal living and for obtaining and processing food (Passingham 1982), but along with their large brains, this adaptation allows primates to interact with and manipulate their environment in unique ways. With the use of their hands, humans have become proficient at manufacturing tools to solve problems and construct limitless structures and apparatuses, and there are many reports showing that NHPs are also capable of making and using tools to hunt, forage, or solve simple tasks. Chimpanzees fish for termites by carefully selecting twigs and stripping off their leaves so they can fit through small termite holes (Boesch and Boesch 1990). Individual chimpanzees also modify their tools for more effective use, fraying the tips of their twigs to provide more surface area for termites to bite. Great apes also show the capacity to use stone tools to forcefully access otherwise inaccessible foods (chimpanzees: Boesch and Boesch 1990; orangutans: van Schaik et al. 1996; gorillas: Breuer et al. 2005), and this ability is present in some species of monkey as well. Some macaques have been known to use stone axes and pounding hammers to process shelled foods (Gumert et al. 2009), and capuchins use carefully selected stone anvils and hammers to break open nuts (Ottoni and Mannu 2001; Fragaszy et al. 2004).

Some NHPs are not only capable of choosing and modifying tools to accomplish a task, they are also capable of using tools to make other tools. Bearded capuchins have been reported selecting small stones to loosen large quartz fragments in conglomerate rock so that the quartz can be used as heavier, more effective hammers in nut cracking (Mannu and Ottoni 2009). This kind of sequential tool use begs the question of the depth of knowledge NHPs use when they are utilizing tools. We have established that NHPs differentiate between different physical properties of objects, but can they use these properties to choose their tools, and does this show that they understand the causal relationship between the physical property and the effectiveness of the tool?

When chimpanzees and human children were shown how to get into an opaque puzzle box to receive a reward, both the chimpanzees and the children could repeat all the steps they had observed in order to receive the treat. However, when a translucent box was introduced, it revealed that most of the steps that had been observed served no functional purpose in retrieving the treat. When chimpanzees and children were again tested on this translucent box, the humans continued to repeat all of the unnecessary steps, but the chimpanzees skipped right to the steps required to access the reward (Horner and Whiten 2005). This result is strong evidence for an NHP's understanding of causality. The human child's inability to complete this task in fewer steps may not necessarily indicate an inability to reason about causality but rather that they may default to overimitating more knowledgeable sources at these early ages.

Other recent experiments show that NHPs are indeed attentive to the physical properties of the tool, selecting tools on the basis of their properties that would best crack open nuts (chimpanzees: Schrauf et al. 2012; capuchins: Visalberghi et al. 2009). This suggests that NHPs can encode the requirements that a tool should meet in order to be effective. We also know that the selection and utilization of tools are acquired and perfected throughout several years in an individual's life with the aid of social observation and trial and error (Pouydebat et al. 2006), and understanding causality is not necessary when repeated exposure to a problem allows an individual to learn what does or does not work in order to obtain a desired outcome. Solving problems with tools on the basis of previous experience does, however, require an excellent memory, which NHPs most certainly have evolved.

Memory

Great apes are able to employ an incredible memory; they show foresight in selecting, transporting, and saving appropriate tools for an apparatus in a different room that they have not seen for up to 14 h (Mulcahy and Call 2006) and replicating novel actions after 24 h (Hopper 2010). Furthermore, both monkeys and apes are able to recall serial lists of symbolic images, and this working memory ability often outperforms that of humans (Inoue and Matsuzawa 2007). This memory for symbols requires that NHPs be capable of a certain amount of mental representation. While no NHP species spontaneously uses symbols, they do have the capacity to associate semantic meaning to an arbitrary symbol. Both apes and monkeys have been able to apply Roman numerals to specific quantities and do simple arithmetic (e.g., chimpanzees: Boysen and Berntson 1989; squirrel monkeys: Olthof et al. 1997).

Numerical Cognition

Usually, instead of counting, animals and human infants mentally represent quantities approximately in an analog format, and many studies show that both children and NHPs share cognitive algorithms for encoding and comparing numerical values and applying simple arithmetic (Cantlon et al. 2009; Feigenson et al. 2004; Gallistel 1989; Meck and Church 1983). Even the brain regions recruited during approximate number representations are present in both humans and NHPs (Dehaene et al. 2003; Nieder 2005), and these cumulative findings indicate that NHPs and humans share the foundation for complex, sophisticated mathematical thought.

Extensive research has shown that NHPs are able to spontaneously represent numerosities of up to 9, at least on an ordinal scale (e.g., chimpanzees: Boysen and Berntson 1989; squirrel monkeys: Olthof et al. 1997; rhesus macaques: Brannon and Terrace 1998; cotton-top tamarins: Uller et al. 2001). When one can mentally combine these values to create a new value without having to directly observe that new value, we call this arithmetic. Research has shown that rhesus macaques are capable of simple addition and subtraction. Rhesus are able to spontaneously compute addition operations over large numbers when the ratio difference between the two numbers is significant (Flombaum et al. 2005), they are able to select a larger quantity of food following subtractions of up to three pieces, and they are capable of representing zero and equality when two quantities are contrasted (Sulkowski and Hauser 2001). This simple computational system of analog numerical representation allows for complex algebra, calculus, and differential equations when the human ability to represent numerical values symbolically is incorporated. However, most NHPs only need an understanding of quantity that is sufficient for comparing resources or social group size, and with these selective priorities, more complex computations are a long way off.

Social Cognition

It is not only the physical environment that imposes challenges on an individual; the social environment supplies its own unique set of challenges. There are many benefits to living in social groups, such as increased protection from predators and greater success in intergroup conflicts and resource defense (e.g., Majolo et al. 2008). However, there are potentially significant costs imposed by increased sociality, such as intensified intragroup competition for food and mates and greater exposure to stronger, more aggressive group members. Over evolutionary history, NHPs have developed behavioral strategies and cognitive abilities to cope with the challenges that arise from heightened sociality. In this section, we will address what NHPs know about their social world and what cognitive skills are necessary to coexist in social groups, such as knowledge of third-party relationships, transitive inference of social relationships, and understanding what others see—and possibly what they know as well.

Knowledge of Third-Party Relationships

Ample evidence demonstrates that NHPs are able to discriminate between individual members of their species in visual (e.g., chimpanzees: Parr et al. 2000; Parr and de Waal 1999; rhesus macaques: Parr et al. 2000; Pascalis and Bachevalier 1998), auditory (e.g., vervets: Cheney and Seyfarth 1980, 1982; savannah baboons: Cheney et al. 1995; squirrel monkeys: Snowdon and Cleveland 1980; rhesus macaques: Rendall et al. 1996; Hansen 1976), and olfactory (e.g., ring-tailed lemurs: Scordato and Drea 2007; Palagi and Dapporto 2006) modalities, as well as cross-modally (rhesus macaques: Adachi and Hampton 2011; Sliwa et al. 2010).

Recognition of social associations has been documented across a variety of kin and non-kin relationships. For example, work by Dasser (1988) has shown that in long-tailed macaques, conspecifics are able to identify mother-offspring dyads as compared to unrelated female-infant dyads and to match a photograph of an infant with the corresponding photograph of its mother. Work in wild vervet monkeys found similar results; using a playback method in which an auditory stimulus of an infant vocalization was played in the presence its mother and two additional adult females, Cheney and Seyfarth (1980) found that the mother oriented toward the origin of the vocalization, whereas the adult females oriented toward the mother prior to the mother's movement. Both studies illustrate that group members are aware of the special relationship between mother and infant and are able to differentiate that relationship on the basis of visual and auditory clues.

NHPs that live in social groups have to keep track of both kin and non-kin relationships, especially with respect to the relative ranks of their social partners. The majority of primate species have clearly established dominance hierarchies, which reduce the incidence of physically aggressive encounters between individuals (reviewed in Kaufmann 1983; e.g., gelada baboons: Bergman et al. 2005; olive baboons: Sapolsky 1992). However, having dominance hierarchies is of little use if individuals are unable to recognize and remember specific relationships between conspecifics and other members of their social group. There is ample observational evidence from studies of both captive and wild populations of NHPs illustrating that NHPs are aware of their own social status and respond appropriately when they encounter a conspecific of higher or lower rank (e.g., in rhesus macaques, a subordinate individual fear-grimacing or presenting their hindquarters to a more dominant individual). Additionally, both observational and experimental evidence suggest that some species of primates are able to "eavesdrop" on interactions between others and extrapolate from observed dominance relationships between two individuals to infer their own dominance to an unknown participant in that interaction, as long as their own dominance relationship to one of the partners in the interaction is known (e.g., if A knows it is dominant to B, and A observes B being dominant to C, A is aware that it is dominant to C, without ever having to interact with C directly).

Transitive Inference

The ability to produce appropriate responses to novel pairings of nonadjacent members in an ordered series without having previous experience with those pairings is known as transitive inference and can be applied to a variety of social interactions, not simply dominance interactions (e.g., reputation learning in chimpanzees: Subiaul et al. 2008, kin-biased redirected aggression in vervets: Cheney and Seyfarth 1999, and recruitment of allies in bonnet macaques: Silk 1999). Most evidence for transitive inference comes from observational data (e.g., vervets: Cheney and Seyfarth 1982, 1986; rhesus macaques: Judge 1982; Japanese macaques: Schino et al. 2006; Aureli et al. 1992; baboons: Bergman et al. 2003), but it has been studied experimentally in a few species as well (e.g., ring-tailed and mongoose lemurs: MacLean et al. 2008; capuchins: D'Amato et al. 1985; vervets: Borgeaud et al. 2013; chimpanzees: Gillan 1981). In an experimental study on prosimians, MacLean et al. (2008) found that although both ring-tailed and mongoose lemurs displayed some evidence of transitive inference on a computer task, the ring-tailed lemurs were markedly better than the mongoose lemurs at this task, which the authors speculate is due to their larger social group size and thus more cognitively complex social environment (but see le Roux and Bergman 2011 for evidence in gelada baboons that knowledge of third-party relationships has limits in extremely large social groups).

Gaze Following: Seeing What Others See or Knowing What Others Know?

Having established that NHPs pay attention to social relationships between other individuals, and that some species are capable of inferring additional relationships that they do not observe or experience directly, it is necessary to establish what behavioral clues individuals use to interpret these relationships. Many of the above relationships can be determined by observation of obvious behavioral interactions, such as affiliation (e.g., grooming, mating) or aggression (e.g., biting, chasing, threatening, etc.). However, there is some evidence that NHPs are also capable of picking up on more subtle social cues as well, such as the direction in which an individual is looking and what it looks at.

It has been well established that NHPs pay attention to conspecifics; however, they do not attend to all physical aspects of conspecifics equally. NHPs attend preferentially to the faces of others, paying particular attention to the eye region (Perrett and Mistlin 1990; Keating and Keating 1982); preferential attention to eyes and eye movements increases gradually across development (Ferrari et al. 2000). The primate brain contains neurons that preferentially respond to head orientation and eye gaze (Perrett et al. 1992; Felleman and Van Essen 1991), and the facial musculature of primates is more highly developed than non-primate animals (Huber 1961; Huber 1931). Research on captive NHPs has revealed that many species engage in gaze following (e.g., ring-tailed lemurs: Shepherd and Platt 2008; sooty mangabeys, rhesus macaques, stump-tailed macaques, pig-tailed macaques, chimpanzees: Tomasello et al. 1998), and anecdotal and observational data from wild populations suggest that gaze following occurs in other primate species as well (e.g., long-tailed macaques: de Waal et al. 1976; hamadryas baboons: Kummer 1967; Savannah baboons: Byrne and Whiten 1992; chimpanzees: Plooij 1978).

Gaze following can provide important information about the socio-ecological environment. By paying attention to what others are attending to, individuals can increase their own knowledge by exploiting the information that others have (e.g., food source location, predator detection, potential mates, etc.). For instance, an experiment conducted on free-ranging rhesus macaques found that when subjects were given a choice between stealing a grape from a human experimenter who was looking at the food or an experimenter who was looking away from the food, the monkey chose to steal from the one not gazing at the grape (Flombaum and Santos 2005). This held true whether the eyes alone or eyes and head were oriented away from the food, as well as when the eyes of one experimenter were occluded by a physical barrier (Flombaum and Santos 2005).

Gaze can also be affected by the social context. In a study on wild black-crested macaques, Micheletta and Waller (2012) found that individuals were quicker to follow the gaze of friends (i.e., conspecifics with whom an individual engaged in frequent, affiliative interactions) than non-friends, suggesting that black-crested macaques prefer information-gathering from friends, or that they are more attuned to social cues from friends. In addition to differences in gaze following within a species, there appear to be differences in how gaze is followed between species. For example, Santos and Hauser (1999) found that cotton-top tamarins cannot use eye gaze alone to predict future behavior of a human experimenter but rather need head and eye orientation in order to predict future action accurately. Using a combination of gaze and gesture, MacLean and Hare (2011) found that chimpanzees and bonobos were more likely to look at a target object that had been looked at and gesticulated toward previously, as compared to a previously ignored object, suggesting that they are attuned to what others are paying attention to.

It has been suggested that gaze following is a prerequisite for other cognitive abilities such as joint attention, predicting the actions of others, social learning, and imitation (Emery et al. 1997). However, what gaze following actually means in terms of cognitive understanding is still up for debate, with some arguing that gaze following is a simple reaction to cues such as body or eye orientation (e.g., Bering and Povinelli 2003), and others contending that gaze following implies that individuals are aware of what others see and possibly what others know (e.g., chimpanzees: Hare et al. 2001; debate reviewed in Seed and Tomasello 2010, as well as Povinelli and Vonk 2004). In part, this debate results from ambiguous data. For example, Povinelli and Eddy (1996a) presented young chimpanzees with two humans, one of whom could see the chimpanzee giving a food begging gesture and one of whom could not. They varied the means by which one human was unable to see the begging gesture (i.e., gazing away from the chimpanzee, eyes blindfolded, head covered by a bucket, or back turned), and the chimpanzees begged indiscriminately from both humans in all conditions except for one—when the human experimenters back was turned. Povinelli and Eddy (1996a) took this as evidence that chimpanzees can be aware of what others can and cannot see but only when the visual clue was very obvious. Regardless of whether behavioral responses to gaze or even head orientation are the product of learned associations or a deeper understanding, several species of NHPs appear to be able to use gaze to gather information about or exploit the attentional state of another. However, the debate about what NHPs know regarding the mental states of others is ongoing and stretches from the domains of gaze following to the understanding (or lack thereof) of intentional versus accidental action and to the major debate over whether NHPs possess a Theory of Mind.

Understanding the Mental States of Others: Theory of Mind

Theory of Mind refers to the ability to attribute mental states, such as perception, beliefs, intentions, desires, and knowledge, to others. Premack and Woodruff (1978) investigated if one adult chimpanzee was able to infer the goals and intentions of a human actor, and although they originally came to the conclusion that their subject understood what the human actor was attempting to do (i.e., the actor's original goal), subsequent research over the next three decades suggested that Premack and Woodruff's result was problematic due to methodological issues and that, in fact, chimpanzees and other NHPs were unable to understand the mental states of others (reviewed in Tomasello and Call 1997).

However, the past decade has brought a sea change in how many primatologists and comparative psychologists study Theory of Mind, and these new methods have produced evidence which some believe suggests that a few species of

NHPs do possess some components of a Theory of Mind (e.g., Seed and Tomasello 2010; Tomasello et al. 2003). Perhaps the most significant factor to spur this change was the realization that previous tests of Theory of Mind in NHPs were not utilizing the natural behaviors and environmental conditions of each species (e.g., Hare and Tomasello 2004).

As complex cognition is hypothesized to have evolved to solve problems encountered in the socio-ecological environment, researchers realized that they could exploit natural tendencies in behavior to gain insight into the NHP's mind, and that the negative results that had been found thus far might be a consequence of trying to elicit responses that were contrary to a species' natural behaviors (Hare and Tomasello 2004). For example, many studies had found that chimpanzees did not utilize gaze cues from human experimenters to select a container that had food, and that they would choose indiscriminately between two containers (e.g., Itakura et al. 1999; Call et al. 1998; Tomasello et al. 1997; Povinelli and Eddy 1996b). This was taken as evidence that chimpanzees did not understand that gaze is an indicator of what others were attending to or interested in and thus that chimpanzees were not able to reason about the knowledge states of others.

However, chimpanzees do not typically cooperate to acquire food, and Hare et al. (2000) hypothesized that they would perform better when tested in a competitive situation. Putting a dominant and subordinate individual in the same testing area, in sight of each other but separated by two clear barriers, Hare et al. (2000) placed two pieces of food in the arena; one food was in an open area where both animals could see it, and the other was placed behind a visual barrier where only the subordinate had visual access to it. Hare et al. (2000) found that when the subordinate was released into the food arena prior to the dominant individual, the subordinate went behind the visual barrier and ate the piece of food that the dominant could not see. Using an almost identical task, Hare et al. (2003) did not find evidence that capuchin monkeys are cognizant of what another individual sees. Hare et al. (2001) suggested that this demonstrates that chimpanzees, but not capuchins, are aware of what others can and cannot see, and that in this competitive context, they will exploit that information to gain food rewards while avoiding negative social consequences. In addition to demonstrating that chimpanzees can use gaze to infer what others can see, and perhaps what they know (Hare et al. 2001), this finding also illustrates the importance of conducting tests in an ecologically relevant manner, which allows animals to behave in a natural way (e.g., Bräuer et al. 2007; Hare and Tomasello 2004).

These new evidence and methods have not convinced everyone, however, and the ensuing controversy and debate have caused a schism in the research community on Theory of Mind in nonhumans. In contrast to those who accept that the current methods allow researchers to investigate what one individual knows about the mind of another (as reviewed in Seed and Tomasello 2010), there are those who believe that researchers have yet to devise a test that accurately distinguishes between subjects who only use behavioral cues and learned associations and subjects who use a combination of behavioral cues/associations as well as inferring the mental states of others (Povinelli and Vonk 2004). Although those who espouse the latter position do not deny the potential existence of a Theory of Mind in NHPs (but see Penn and Povinelli 2007), they find the current data inadequate in distinguishing between whether individuals are only paying attention to behavior or whether they are inferring something about the mental states of others in addition to reading behavioral clues (e.g., Povinelli and Vonk 2004; Karin-D'Arcy and Povinelli 2002).

It has been suggested that the best way to rectify the current crisis is to use versions of tasks that test false beliefs in verbal subjects (e.g., the Sally-Anne or Smarties task often used to study Theory of Mind in human children) that are adapted for use in non- or preverbal individuals (Lurz and Krachun 2011). Thus far, few studies have used false-belief tasks to investigate Theory of Mind in nonhuman primates (but see Call and Tomasello 1999). One study in rhesus macaques, using a violation-of-expectancy paradigm commonly used in studies of human infants (Spelke

1985), found that subjects exhibited surprise (by looking longer) when a human experimenter looked in the wrong location for an object when that experimenter knew where the object was, by virtue of having seen it placed there earlier (Marticorena et al. 2011). However, the rhesus macaques did not appear to make any prediction (i.e., they looked equally at all possible locations) about where the experimenter would look when the object was moved from its original location without the experimenter's knowledge (Marticorena et al. 2011). This suggests that although rhesus macaques can predict what an individual will do when that individual has a true belief, they cannot predict behavior when an individual has a false belief (Marticorena et al. 2011). Importantly, none of the species that have been tested thus far have demonstrated an understanding of false beliefs (e.g., Krachun et al. 2009a, b; Kaminski et al. 2008; Hare et al. 2001; Call and Tomasello 1999).

Keeping in mind the possibility that what appears to be a cognitively complex skill could in fact be attributed to learned associations between behaviors and subsequent actions, we will review the literature which suggests that NHPs might have an understanding of the mental states of both themselves and others.

Concept of Self

Before scientists became interested in the question of whether NHPs had a concept of the minds of others, they were interested in whether individuals had a concept of self (e.g., Gallup 1970). The relationship between understanding that one has his/her own mind and being able to attribute mentality to others is not well understood, even in humans (e.g., Happé 2003). However, it is almost certain that a concept of self is necessary for a fully developed Theory of Mind (Gallup 1982).

Gallup (1970) observed that when chimpanzees were placed in front of a mirror, they initially reacted as though their reflection were another chimpanzee and acted socially towards it. However, after a few days of mirror exposure, the chimpanzees began to use the mirror to act toward themselves, grooming body parts that were not visible without the aid of the mirror (Gallup 1970). In light of these observations and with the hypothesis that in order to display self-recognition, an individual must have a concept of self, Gallup developed the mark test to investigate whether non- or preverbal individuals displayed mirror self-recognition (MSR) (Gallup 1977). In the mark test, while a subject is unconscious or distracted, a mark in a conspicuous color is painted onto a part of the body that the individual cannot feel or see without a mirror (usually its forehead). Once awake, the subject is placed in front of a mirror and the number of touches directed to the mark are counted, as are the number of touches that an individual makes to other parts of their body. This test has been applied to a wide variety of species, and at present it is established that humans and most great apes pass the mark test by touching the mark more frequently in front of the mirror than when no mirror is present (chimpanzees: de Veer et al. 2002; Gallup 1977), but monkeys do not pass this test (reviewed in Anderson and Gallup 2011; Povinelli 1987) and show persistent social responses to mirrors over prolonged exposure.

However, as with most areas of primate intelligence, the line in the sand between monkeys and apes is under dispute, with some suggesting that an inability to pass the mark test does not mean that monkeys do not have the ability for self-recognition (e.g., Seyfarth and Cheney 2000; Heyes 1994). There is some evidence that monkeys do in fact show self-directed behaviors when in front of a mirror that do not take place in its absence (e.g., rhesus macaques: Rajala et al. 2010) and that they are capable of equating videos of parts of their own bodies with themselves (e.g., capuchins: Anderson et al. 2009; Japanese macaques: Iriki et al. 2001), and this indicates that at least precursors to self-recognition may be present in some monkey species (e.g., pygmy marmosets: Eglash and Snowdon 1983). Moreover, one species of great apes does not pass the mark test—gorillas—suggesting that the evolution of self-recognition does not divide monkeys and apes, but that performance on the mark test is affected by other factors such as a natural inclination not to look directly at others that may diminish an individual's willingness to participate in the test (reviewed in Suarez and Gallup 1981). Drawing from inconsistencies in the data on MSR, it has been suggested that the visual modality might not always be the best test for self-recognition, and that some species might perform better in an auditory or olfactory modality or even simply by testing for an understanding of oneself versus others in a more social setting (Seyfarth and Cheney 2000).

Inferring Intentions

Some of the first work to investigate what NHPs understand about the mental states of others came from studies asking whether primates could predict the goals of others. This cognitive ability would provide benefits for individuals in their natural environment, enabling them to predict not only what others will do in highly prescribed situations but also in novel ones (Call and Tomasello 2008). What NHPs understand about the goals of others can be tested in multiple ways.

First, we can ask whether NHPs are able to discriminate between intentional and accidental actions. For example, Povinelli et al. (1998) tested whether chimpanzees preferred to receive food from an experimenter who had previously been clumsy (e.g., accidentally spilling juice when trying to hand it to the chimp) or an experimenter who intentionally spilled the juice instead of handing it. Although the results were not strong, they found a general preference for the clumsy experimenter as compared to the one who intentionally did not deliver the juice (Povinelli et al. 1998), suggesting that chimpanzees might be able to discriminate between accidents and intentional actions (see Call and Tomasello 1998 for evidence of this ability in orangutans).

Second, this inspires a similar question of whether NHPs are able to differentiate between individuals who are unable to help and those who are unwilling to help. To test this, Call et al. (2004) had a human experimenter give food to a chimpanzee until the experimenter became either unwilling or unable to give food any longer. Call et al. (2004) found that the subject's rate of frustration behaviors was much higher when the experimenter was unwilling to dispense food than when the experiment was unable. Additionally, the subjects left the testing area and terminated the session earlier in the unwilling condition, again suggesting that chimpanzees can discriminate between humans who will not help them because they cannot and those who will not help because they do not want to (Call et al. 2004). Although few monkey species have been tested on their understanding of intentions, capuchins have also shown the ability to discriminate between an experimenter that is unwilling and one that is unable (Phillips et al. 2009).

Third, we can test attentiveness to intentionality by observing what types of helping behavior NHPs offer when they observe an experimenter struggling to achieve a particular goal. Using this method, Warneken and Tomasello (2006) had a human experimenter accidentally or intentionally drop an object they were carrying and observed whether or not chimpanzees would retrieve the object and return it to the experimenter; they found that chimpanzees would help in the accidental but not the intentional condition. Warneken et al. (2007) also employed a helping paradigm using two chimpanzees, where one could assist the other in their goal of opening a door, and found similar results. Capuchins also demonstrated an understanding of intentions, by offering help to a human experimenter (Barnes et al. 2008), but in contrast to the chimpanzees (Warneken et al. 2007), the capuchins would not assist the experimenter if there was a cost to themselves. This illustrates an interesting species difference that should be taken into consideration when testing other primates, as other species may abandon helping behavior in certain contexts, resulting in null test results, even if they are mentally capable of understanding the experimenter's goal.

Fourth, we can test NHPs understanding of intentionality by testing their imitation of rational or irrational behavior. In an experiment using methods co-opted from studies of human infants (Gergely et al. 2002), enculturated chimpanzees were exposed to a human experimenter carrying out an action in an unusual way because the experimenter was physically constrained (e.g., the experimenter opens a door with his foot instead of his hands because his hands are occupied carrying an object: Buttelmann et al. 2007) and to an experimenter carrying out the same unusual action even though there was no reasonable constraint. Buttelmann et al. (2007) found that the chimpanzees were more likely to imitate the experimenter's unusual actions (using their foot) to achieve the goal of opening the door after observing the unconstrained condition, but they would simply open the door with their hands following observation of the constrained condition. This suggests that the chimpanzees were aware of the goal of the experimenter and were able to conclude that the experimenter in the constrained condition was only using his/her feet because they could not use their hands, but the experimenter in the unconstrained condition was using their feet for some other reason (see Buttelmann et al. 2012 for similar results using a novel method).

Interestingly, Buttelmann et al. (2007) did not find the same result with non-enculturated chimpanzees (see also Tennie et al. 2012), suggesting that imitation is a learned behavior in chimpanzees. Previous studies have also suggested that chimpanzees and the other great apes are actually better emulators (carrying out an intended goal without carrying out the specific actions demonstrated in achieving that goal) than imitators (mimicking specific actions step by step), as compared to human children who are prone to overimitation (e.g., Tennie et al. 2006, 2010; Horner and Whiten 2005; but see Whiten et al. 2009 for evidence of imitation in wild chimpanzees). Compared to great apes, monkeys have been studied much less, yet there is some evidence that capuchin monkeys are able to imitate the actions of others to achieve a goal (van de Waal and Whiten 2012). However, observational data from wild and free-ranging populations of monkeys typically suggest that imitation is infrequent and rarely utilized in natural contexts (reviewed in Visalberghi and Fragaszy 1990).

Together, these methods and studies indicate that NHPs are capable of understanding the intentions of others, and they can discriminate between an intentional and an accidental action. Still, this is not well studied outside of great apes, and even within great apes, most research has been conducted on chimpanzees. More work is necessary to gain insight into whether other NHPs have a concept of intentionality, both using experimental methods as well as through observation of wild populations.

Deception

Observation of wild and free-ranging populations has actually produced some of the most compelling evidence that NHPs understand the mental states of others, especially in the domain of deception. For example, observational research has revealed evidence of tactical deception that emerges in multiple modalities. NHPs will stay quiet and withhold vocalizations during certain situations (e.g., sneaky mating, competing for food, territory defense in wild chimpanzees: Watts and Mitani 2001; Wilson et al. 2001), especially when alerting others to their location could result in aggression (e.g., evidence of punishment for withholding food calls in rhesus macaques: Hauser 1992). Le Roux et al. (2013) found evidence for tactical deception in gelada baboons, where males and females were less likely to vocalize during extra-pair copulations and were more likely to engage in this sneaky mating when the female's primary mate was further away, thus avoiding an aggressive punishment.

Experimental studies of deception can also shed light on NHPs understanding of the perspective and knowledge of others. For example, both rhesus macaques (Santos et al. 2006) and chimpanzees (Melis et al. 2006) choose to steal food from a silent (nonfunctional bells attached) as opposed to noisy (functional bells attached) container when there is a food competitor present (but see Bräuer et al. 2008 for a contradictory result in chimpanzees). Additionally, Hare et al. (2006) found that when chimpanzees were put in a competitive situation with a human experimenter over a contested piece of food, the chimpanzees would manipulate the visual perception of the experimenter to make their approach to the food item less obvious (i.e., by hiding behind an occluder) and thereby hiding their intention to grab the food from the experimenter. Although there is evidence that rhesus macaques are able to deceive others in visual (Flombaum and Santos 2005) and auditory (Santos et al. 2006) domains, not all species of monkeys appear able to do so (e.g., long-tailed macaques: Kummer et al. 1996).

Returning to the idea that challenges in the social environment spurred the evolution of increased cognition and larger brains, multiple relationships between deception and various species of NHPs have been noted to better understand its evolution in primates. In conducting a meta-analysis of previously published papers, Byrne and Whiten (1992) found that deception is not found across all species and clades of NHPs equally; instead, deception is found most frequently in great apes and infrequently, if at all, in prosimians. Additionally, Byrne and Corp (2004) found a relationship between neocortical volume and use of deception across the Primate order. In order to understand this relationship more fully, additional studies on previously untested species of NHPs are necessary.

A Comparative Approach

A call for study of additional species is common and is often repeated in the articles and studies described above. Within the order Primates, there is a remarkable diversity of species, and extensive studies of many primate species have shown that these species exist in a wide variety of habitats and exhibit many types of mating and social systems as well as behaviors. While much research thus far has described socio-ecological differences between species, far fewer studies have examined species-specific differences in cognition. As primates are considered one of the most cognitively complex taxa, with the evolution of primate intelligence hypothesized to be driven by their high degree of sociality (Dunbar 1998), the utility of a comparative approach to the study of cognition cannot be overstated (MacLean et al. 2012), especially as differences in cognition could very well be both shaped by and have consequences for social relationships across a given species' lifespan.

Despite the acknowledged utility for a truly comparative approach to studies of intelligence, few species of primates have been studied as thoroughly as the great apes, with a particular attention on chimpanzees and humans. We expect vast differences in the cognitive abilities of great apes and humans, as the human brain size is three times that of the other great apes and humans show a much larger behavioral repertoire, but previous research has shown that humans, at least early in development are not any more intelligent in general than great apes (reviewed in Tomasello and Hermann 2010). This was best demonstrated through the results of a battery of tests administered to chimpanzees, orangutans, and 2-year-old humans (Hermann et al. 2007) in which humans' socio-cognitive abilities (e.g., reading intentions, communicating, and learning socially) were greatly enhanced relative to the other great apes, but all three species were equally able to reason about the physical world (e.g., understanding causality, space, and quantity) (see also Hermann et al. 2010).

Tomasello et al. (2005) hypothesized that the crucial difference between humans and nonhuman primates is humans' ability to reason, understand, learn from, and act effectively in their unique social environments in a concept we call culture. But while humans are undeniably special and intelligent in a way that is not seen in nonhuman species, it is perhaps premature to conclude that the crucial difference between humans and NHPs is their ability to reason and know about the social world. Many species of NHPs live in extraordinarily complex social groups (e.g., gelada baboons: Snyder-Mackler et al. 2011), even larger and more complex than the social groups early hominids were thought to live in. As shown by Hare et al. (2001), sometimes the question might simply be investigated in an unsuitable way, making it unintelligible to the species being tested. As we are humans, we better understand the best way to ask other humans what they think and know about themselves and others, but more thought and consideration, coupled with an understanding of the social and ecological context a species inhabits, is necessary to devise studies that are relevant and sensible to NHPs and allow us to distinguish between the true mental states of NHPs and simple behavioral responses. The past decade has brought remarkable advances in our understanding of primate intelligence (Whiten 2013), and by utilizing a comparative approach and devising clever, species-specific cognitive tests, the next decade promises to be just as productive.

References

Adachi, I., & Hampton, R. R. (2011). Rhesus monkeys see who they hear: Spontaneous cross-modal memory for familiar conspecifics. *PLoS One, 6*, 1–8.

Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution. *Current Anthropology, 36*, 199–221.

Anderson, J. R., & Gallup, G. G., Jr. (2011). Which primates recognize themselves in mirrors? *PLoS Biology, 9*, 1–3.

Anderson, J. R., Kuroshima, H., Paukner, A., & Fujita, K. (2009). Capuchin monkeys (Cebus apella) respond to video images of themselves. *Animal Cognition, 12*, 55–62.

Aureli, F., Cozzolino, R., Cordischi, C., & Scucchi, S. (1992). Kin-oriented redirection among Japanese macaques—An expression of a revenge system. *Animal Behaviour, 44*, 283–291.

Barnes, J. L., Hill, T., Langer, M., Martinez, M., & Santos, L. R. (2008). Helping behaviour and regard for others in capuchin monkeys (Cebus apella). *Biology Letters, 4*, 638–640.

Barrett, L., Henzi, P., & Rendell, D. (2007). Social brains, simple minds: Does social complexity really require cognitive complexity? *Philosophical Transactions of the Royal Society B, 362*, 561–575.

Bergman, T. J., Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science, 302*, 1234–1236.

Bergman, T. J., Beehner, J. C., Cheney, D. L., Seyfarth, R. M., & Whitten, P. L. (2005). Correlates of stress in free-ranging male chacma baboons, Papio hamadryas ursinus. *Animal Behaviour, 70*, 703–713.

Bering, J. M., & Povinelli, D. J. (2003). Comparing cognitive development. In D. Maestripieri (Ed.), *Primate psychology* (pp. 205–233). Cambridge, MA: Harvard University Press.

Boesch, C., & Boesch, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatologica, 54*, 86–99.

Borgeaud, C., van de Waal, E., & Bshary, R. (2013). Third-party ranks knowledge in wild vervet monkeys (Chlorocebus aethiops pygerythrus). *PLoS One, 8*, 1–4.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (Pan troglodytes). *Journal of Comparative Psychology, 103*, 23–31.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science, 282*, 746–749.

Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees really know what others can see in a competitive situation. *Animal Cognition, 10*, 439–448.

Bräuer, J., Call, J., & Tomasello, M. (2008). Chimpanzees do not take into account what others can hear in a competitive situation. *Animal Cognition, 11*, 175–178.

Breuer, T., Ndoundou-Hockemba, M., & Fishlock, V. (2005). First observation of tool use in wild gorillas. *PLoS Biology, 3*, 2041–2043.

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental Science, 10*, F31–F38.

Buttelmann, D., Schütte, S., Carpenter, M., Call, J., & Tomasello, M. (2012). Great apes infer others' goals based on context. *Animal Cognition, 15*, 1037–1053.

Byrne, R. W., & Corp, N. (2004). Neocortex predicts deception rate in primates. *Proceedings of the Royal Society of London: Series B, 271*, 1693–1699.

Byrne, R. W., & Whiten, A. (1992). Cognitive evolution in primates: Evidence from tactical deception. *Man, 27*, 609–627.

Cacchione, T., & Horst, K. (2004). Recognizing impossible object relations: Intuitions about support in chimpanzees (Pan troglodytes). *Journal of Comparative Psychology, 118*, 140–148.

Call, J., & Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology, 112*, 192–206.

Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development, 70*(2), 381–395.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*, 187–192.

Call, J., Hare, B., & Tomasello, M. (1998). Chimpanzee gaze following in an object-choice task. *Animal Cognition, 1*, 89–99.

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). 'Unwilling' versus 'unable': Chimpanzees' understanding of human intentional action. *Developmental Science, 7*, 488–498.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13*, 83–91.

Cheney, D. L., & Seyfarth, R. M. (1982). Recognition of individuals within and between groups of free-ranging vervet monkeys. *American Zoologist, 22*, 519–529.

Cheney, D. L., & Seyfarth, R. M. (1986). The recognition of social alliances among vervet monkeys. *Animal Behaviour, 34*, 1722–1731.

Cheney, D. L., & Seyfarth, R. M. (1999). Recognition of other individuals' social relationships by female baboons. *Animal Behaviour, 58*, 67–75.

Clutton-Brock, T. H., & Harvey, P. (1980). Primates, brains, and ecology. *Journal of Zoology, 190*, 309–323.

D'Amato, M. R., Salmon, P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (Cebus paella) and pigeons (Columba livia). *Journal of the Experimental Analysis of Behavior, 44*, 35–47.

Dasser, V. (1988). A social concept in Java monkeys. *Animal Behaviour, 36*, 225–230.

de Bois, S. T., & Novak, M. A. (1994). Object permanence in rhesus monkeys (Macaca mulatta). *Journal of Comparative Psychology, 108*, 318–327.

de Bois, S. T., Novak, M. A., & Bond, M. (1998). Object permanence in orangutans (Pongo pygmaeus) and squirrel monkeys (Saimiri sciureus). *Journal of Comparative Psychology, 112*, 137–152.

de Veer, M. W., Gallup, G. G., Theall, L. A., van den Bos, R., & Povinelli, D. J. (2002). An 8-year longitudinal study of mirror self-recognition in chimpanzees (Pan troglodytes). *Neuropsychologia, 1493*, 1–6.

de Waal, F. B. M., van Hooff, J. A. R. A. M., & Netto, W. J. (1976). An ethological analysis of types of agonistic interaction in a captive group of Java-monkeys (Macaca fascicularis). *Primates, 17*, 257–290.

Deaner, R. O., van Schaik, C. P., & Johnson, V. (2006). Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evolutionary Psychology, 4*, 149–196.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology, 6*, 178–190.

Dunbar, R. I. M., & Shultz, S. (2007). Evolution in the social brain. *Science, 317*, 1344–1347.

Eglash, A. R., & Snowdon, C. T. (1983). Mirror-image responses in pygmy marmosets (Cebuella pygmaea). *American Journal of Primatology, 5*, 211–219.

Emery, N. J., Lorincz, E. N., Perrett, D. I., Oram, M. W., & Baker, C. I. (1997). Gaze following and joint attention in rhesus monkeys (Macaca mulatta). *Journal of Comparative Psychology, 111*, 286–293.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.

Ferrari, P. F., Kohler, E., Fogassi, L., & Gallese, V. (2000). The ability to follow eye gaze and its emergence during development in macaque monkeys. *Proceedings of the National Academy of Science, 97*, 13997–14002.

Fish, J. L., & Lookwood, C. A. (2003). Dietary constraints on encephalization in primates. *American Journal of Physial Anthropology, 120*, 171–181.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology, 15*, 447–452.

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (Macaca mulatta) spontaneously compute addition operations over large numbers. *Cognition, 97*, 315–325.

Fragaszy, D., Izar, P., Visalberghi, E., Ottoni, E. B., & de Oliveira, M. G. (2004). Wild capuchin monkeys (Cebus libidinosus) use anvils and stone pounding tools. *American Journal of Primatology, 64*, 359–366.

Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Annual Review of Psychology, 40*, 155–189.

Gallup, G. G., Jr. (1970). Chimpanzees: Self-recognition. *Science, 167*, 86–87.

Gallup, G. G., Jr. (1977). Self-recognition in primates: A comparative approach to the bidirectional properties of consciousness. *American Psychologist, 32*, 329–338.

Gallup, G. G., Jr. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology, 2*, 237–248.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature, 415*, 755.

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes, 7*, 150–164.

Grueter, C. C., Bissonnette, A., Isler, K., & Van Schaik, C. P. (2013). Grooming and group cohesion in primates: Implications for the evolution of language. *Evolution and Human Behavior, 34*, 61–68.

Gumert, M. D., Kluck, M., & Malaivijitnond, S. (2009). The physical characteristics and usage patterns of stone axe and pounding hammers used by long-tailed macaques in the Andaman Sea region of Thailand. *American Journal of Primatology, 71*, 594–608.

Hansen, E. W. (1976). Selective responding by recently separated juvenile rhesus monkeys to the calls of their mother. *Developmental Psychobiology, 9*, 83–88.

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology, 18*, 370–372.

Happé, F. (2003). Theory of mind and the self. *Annals of the New York Academy of Sciences, 1001*, 134–144.

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skillful in competitive than in cooperative cognitive tasks. *Animal Behaviour, 68*, 571–581.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour, 59*, 771–785.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour, 6*, 139–151.

Hare, B., Addessi, E., Call, J., Tomasello, M., & Visalberghi, E. (2003). Do capuchin monkeys, Cebus apella, know what conspecifics do and do not see? *Animal Behaviour, 65*, 131–142.

Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition, 10*, 495–514.

Hauser, M. D. (1992). Costs of deception: Cheaters are punished in rhesus monkeys (Macaca mulatta). *Proceedings of the National Academy of Sciences of the United States of America, 89*, 12137–12139.

Hauser, M. D. (1998). A non-human primate's expectations about object motion and destination: The importance of self- propelled movement and animacy. *Developmental Science, 1*, 31–38.

Hermann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360–1366.

Hermann, E., Hernandez-Lloreda, M. V., Call, J., Hare, B., & Tomasello, M. (2010). The structure of individual differences in the cognitive abilities of children and chimpanzees. *Psychological Science, 21*, 102–110.

Heyes, C. M. (1994). Reflections on self-recognition in primates. *Animal Behaviour, 47*, 909–919.

Holekamp, K. E. (2007). Questioning the social intelligence hypothesis. *Trends in Cognitive Sciences, 11*, 65–69.

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a non-human primate? *Developmental Science, 2*, 35–41.

Hopper, L. M. (2010). Deferred imitation in children and apes: Children imitate after a delay, but can apes ape in a similar fashion? *The Psychologist, 23*, 2–5.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition, 8*, 164–181.

Huber, E. (1931). *Evolution of facial musculature and facial expression*. Baltimore: John Hopkins Press.

Huber, E. (1961). The facial musculature and its innervation. In C. G. Hartman & W. L. Straus (Eds.), *Anatomy of the rhesus monkey (Macaca mulatta)* (pp. 176–188). New York: Hafner.

Humphrey, N. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge, UK: Cambridge University Press.

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology, 17*, R1004–R1005.

Iriki, A., Tanaka, M., Obayashi, S., & Iwamura, Y. (2001). Self-images in the video monitor coded by monkey intraparietal neurons. *Neuroscience Research, 40*, 163–173.

Itakura, S., Agnetta, B., Hare, B., & Tomasello, M. (1999). Chimpanzees use human and conspecific social cues to locate hidden food. *Developmental Science, 4*, 448–456.

Judge, P. (1982). Redirection of aggression based on kinship in a captive group of pigtail macaques. *International Journal of Primatology, 3*, 301.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition, 109*, 224–234.

Karin-D'Arcy, R., & Povinelli, D. J. (2002). Do chimpanzees know what each other see? A closer look. *International Journal of Comparative Psychology, 15*, 21–54.

Kaufmann, J. H. (1983). On the definitions and functions of dominance and territoriality. *Biological Reviews, 58*, 1–20.

Keating, C. F., & Keating, E. G. (1982). Visual scan patterns of rhesus monkeys viewing faces. *Perception, 11*, 211–219.

King, B. J. (1986). Extractive foraging and the evolution of primate intelligence. *Human Evolution, 1*, 361–372.

Krachun, C., Call, J., & Tomasello, M. (2009a). Can chimpanzees (Pan troglodytes) discriminate appearance from reality? *Cognition, 112*, 435–450.

Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009b). A competitive nonverbal false belief task for children and apes. *Developmental Science, 12*, 521–535.

Kummer, H. (1967). Tripartite relations in hamadryas baboons. In S. A. Altmann (Ed.), *Social communication among primates* (pp. 63–71). Chicago: The University of Chicago Press.

Kummer, H., Anzenberger, G., & Hemelrijk, C. K. (1996). Hiding and perspective taking in long-tailed macaques (Macaca fascicularis). *Journal of Comparative Psychology, 110*, 97–102.

le Roux, A., & Bergman, T. J. (2011). Indirect rival assessment in a social primate, Theropithecus gelada. *Animal Behaviour, 83*, 249–255.

le Roux, A., Snyder-Mackler, N., Roberts, E. K., Beehner, J. C., &Bergman, T. J. (2013). Evidence for tactical concealment in a wild primate. *Nature Communications*, 4. doi:10.1038/ncomms2468.

Lurz, R. W., & Krachun, C. (2011). How could we know whether nonhuman primates understand others' internal goals and intentions? Solving Povinelli's problem. *Review of Philosophy and Psychology, 2*, 449–481.

MacLean, E. L., & Hare, B. (2011). Bonobos and chimpanzees infer the target of another's attention. *Animal Behaviour, 83*, 345–353.

MacLean, E. L., Merritt, D. J., & Brannon, E. M. (2008). Social complexity predicts transitive reasoning in prosimian primates. *Animal Behaviour, 76*, 479–486.

MacLean, E. L., Matthews, L. J., Hare, B. A., Nunn, C. L., Anderson, R. C., Aureli, F., Brannon, E. M., Call, J., Drea, C. M., Emery, N. J., Haun, D. B. M., Herrmann, E., Jacobs, L. F., Platt, M. L., Rosati, A. G., Sandel, A. A., Schroepfer, K. K., Seed, A. M., Tan, J., van Schaik, C. P., & Wobber, V. (2012). How does cognition evolve? Phylogenetic comparative psychology. *Animal Cognition, 15*, 223–238.

Majolo, B., de Bortoli Vizioli, A., & Schino, G. (2008). Costs and benefits of group living in primates: Group size effects on behaviour and demography. *Animal Behaviour, 76*, 1235–1247.

Mannu, M., & Ottoni, E. B. (2009). The enhanced tool-kit of two groups of monkeys in the Caatinga: Tool making, associative use, and secondary tools. *American Journal of Primatology, 3*, 242–251.

Marticorena, D. C. W., Ruiz, A. M., Mukerji, C., Goddu, A., & Santos, L. R. (2011). Monkeys represent others' knowledge but not their beliefs. *Developmental Science, 14*, 1406–1416.

Martin, R. D. (2013). Growing a large brain. In R. D. Martin (Ed.), *How we do it* (pp. 121–146). New York: Basic Books.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 320–334.

Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology, 120*, 154–162.

Micheletta, J., & Waller, B. M. (2012). Friendship affects gaze following in a tolerant species of macaque, Macaca nigra. *Animal Behaviour, 83*, 459–467.

Milton, K. (1981). Distribution pattern of tropical plant foods as an evolutionary stimulus to primate mental development. *American Anthropologist, 83*, 534–548.

Mireille, M., Bouchard, M. A., Granger, L., & Herscovitch, J. (1976). Piagetian object-permanence in Cebus capucinus, Lagothrica flavicauda and Pan troglodytes. *Animal Behaviour, 24*, 585–588.

Mulcahy, N. J., & Call, J. (2006). Apes save tools for future use. *Science, 312*, 1038–1040.

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience, 6*, 177–190.

Olthof, A., Iden, C. M., & Roberts, W. A. (1997). Judgments of ordinality and summation of number symbols by squirrel monkeys (Saimiri sciureus). *Journal of Experimental Psychology, 23*, 325–339.

Ottoni, E. B., & Mannu, M. (2001). Semi free-ranging Tufted capuchins (Cebus apella) spontaneously use tools to crack open nuts. *International Journal of Primatology, 22*, 347–358.

Palagi, E., & Dapporto, L. (2006). Beyond odor discrimination: Demonstrating individual recognition by scent in Lemur catta. *Chemical Senses, 31*, 437–443.

Parker, S. T., & Gibson, K. R. (1977). Object manipulation, tool use and sensorimotor intelligence as feeding adaptations in cebus monkeys and great apes. *Journal of Human Evolution, 6*, 623–641.

Parker, S. T., & Gibson, K. R. (1979). A developmental model for the evolution of language and intelligence in early hominids. *Behavioral and Brain Sciences, 2*, 367–408.

Parr, L. A., & de Waal, F. B. M. (1999). Visual kin recognition in chimpanzees. *Nature, 399*, 647–648.

Parr, L. A., Winslow, J. T., Hopkins, W. D., & de Waal, F. B. M. (2000). Recognizing facial cues: Individual discrimination by chimpanzees (Pan troglodytes) and rhesus monkeys (Macaca mulatta). *Journal of Comparative Psychology, 114*, 47–60.

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioural Processes, 43*, 87–96.

Passingham, R. E. (1982). *The human primate*. San Francisco: W.H. Freeman.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362*, 731–744.

Perrett, D. I., & Mistlin, A. J. (1990). Perception of facial characteristics by monkeys. In W. C. Stebbins & M. A. Berkley (Eds.), *Comparative perception* (pp. 187–215). New York: Wiley.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London B, 335*, 23–30.

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). 'Unwilling' versus 'unable': Capuchin monkeys' (Cebus apella) understanding of human intentional action. *Developmental Science, 12*, 938–945.

Platt, M. L., Brannon, E. M., Briese, T. A., & French, J. A. (1996). Differences in feeding ecology predict differences in performance between golden lion tamarins (Leontopithecus rosalia) and Wied's marmosets (Callithrix kuhli) on spatial and visual memory tasks. *Animal Learning and Behavior, 24*, 384–393.

Plooij, F. X. (1978). Some basics traits of language in wild chimpanzees? In A. Lock (Ed.), *Action, gesture and symbol* (pp. 111–131). London: Academic.

Pouydebat, E., Gorce, P., Bels, V., & Coppens, Y. (2006). Substrate optimization in nut cracking by capuchin monkeys (Cebus apella). *American Journal of Primatology, 68*, 1017–1024.

Povinelli, D. J. (1987). Monkeys, apes, mirror and minds: The evolution of self-awareness in primates. *Human Evolution, 2*, 493–509.

Povinelli, D. J., & Eddy, T. J. (1996a). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development, 61*, 1–152.

Povinelli, D. J., & Eddy, T. J. (1996b). Chimpanzees: Joint visual attention. *Psychological Science, 7*, 129–135.

Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind and Language, 19*, 1–28.

Povinelli, D. J., Perilloux, H. K., Reaux, J. E., & Bierschwale, D. T. (1998). Young and juvenile chimpanzees' (Pan troglodytes) reactions to intentional versus accidental and inadvertent actions. *Behavioural Processes, 42*, 205–218.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515–526.

Rajala, A. Z., Reininger, K. R., Lancaster, K. M., & Populin, L. C. (2010). Rhesus monkeys (Macaca mulatta) do recognize themselves in the mirror:

Implications for the evolution of self-recognition. *PLoS One, 5*, 1–8.

Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences, 99*, 4436–4441.

Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society B, 366*, 1017–1027.

Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour, 51*, 1007–1015.

Santos, L. R. (2004). 'Core Knowledges': A dissociation between spatiotemporal knowledge and contact-mechanics in a non-human primate? *Developmental Science, 7*, 167–174.

Santos, L. R., & Hauser, M. D. (1999). How monkeys see the eyes: Cotton-top tamarins' reaction to changes in visual attention and action. *Animal Cognition, 2*, 131–139.

Santos, L. R., & Hauser, M. D. (2002). A non-human primate's understanding of solidity: Dissociations between seeing and acting. *Developmental Science, 5*, F1–F7.

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys (Macaca mulatta) know what others can and cannot hear. *Animal Behaviour, 71*, 1175–1181.

Sapolsky, R. M. (1992). Cortisol concentrations and the social significance of rank instability among wild baboons. *Psychoneuroendocrinology, 17*, 701–709.

Schino, G., Tiddi, B., & Polizzi Di Sorrentino, E. (2006). Simultaneous classification by rank and kinship in Japanese macaques. *Animal Behaviour, 71*, 1069–1074.

Schrauf, C., Call, J., Fuwa, K., & Hirata, S. (2012). Do chimpanzees use weight to select hammer tools? *PLoS One, 7*, 1–12.

Scordato, E. S., & Drea, C. M. (2007). Scents and sensibility: Information content of olfactory signals in the ringtailed lemur (Lemur catta). *Animal Behaviour, 73*, 301–314.

Seed, A., & Tomasello, M. (2010). Primate cognition. *Topics in Cognitive Science, 2*, 407–419.

Seyfarth, R. M., & Cheney, D. L. (2000). Social awareness in monkeys. *American Zoologist, 40*, 902–909.

Shepherd, S. V., & Platt, M. L. (2008). Spontaneous social orienting and gaze following in ringtailed lemurs (Lemur catta). *Animal Cognition, 11*, 13–20.

Silk, J. B. (1999). Male bonnet macaques use information about third-party rank relationships to recruit allies. *Animal Behaviour, 58*, 45–51.

Sliwa, J., Duhamel, J., Pascalis, O., & Wirth, S. (2010). Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proceedings of the National Academy of Sciences, 108*, 1735–1740.

Snyder-Mackler, N., Beehner, J. C., & Bergman, T. J. (2011). Defining higher levels in the multilevel societies of Geladas (Theropithecus gelada). *International Journal of Primatology, 33*, 1054–1068.

Spelke, E. S. (1985). Preferential-looking methods as tools for the study of cognition in infancy. In G. Gottlieb & N. A. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 323–363). Westport: Ablex.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*, 29–56.

Suarez, S. D., & Gallup, G. G., Jr. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution, 10*, 175–188.

Subiaul, F., Vonk, J., Okamoto-Barth, S., & Barth, J. (2008). Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Animal Cognition, 11*, 611–623.

Sulkowski, G. M., & Hauser, M. D. (2001). Can rhesus monkeys spontaneously subtract? *Cognition, 79*, 239–262.

Tennie, C., Call, J., & Tomasello, M. (2006). Push or pull: Imitation vs. emulation in great apes and human children. *Ethology, 112*, 1159–1169.

Tennie, C., Call, J., & Tomasello, M. (2010). Evidence for emulation in chimpanzees in social settings using the floating peanut task. *PLoS One, 5*, 1–9.

Tennie, C., Call, J., & Tomasello, M. (2012). Untrained chimpanzees (Pan troglodytes schweinfurthii) fail to imitate novel actions. *PLoS One, 7*, 1–19.

Tomasello, M., & Call, J. (1997). *Primate cognition*. New York: Oxford University Press.

Tomasello, M., & Hermann, E. (2010). Ape and human cognition: What's the difference? *Current Directions in Psychological Science, 19*, 3–8.

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour, 55*, 1063–1069.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends in Cognitive Sciences, 7*, 153–156.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–735.

Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (Saguinus oedipus). *Journal of Comparative Psychology, 115*, 248–257.

van de Waal, E., & Whiten, A. (2012). Spontaneous emergence, imitation and spread of alternative foraging techniques among groups of vervet monkeys. *PLoS One, 7*, 1–7.

van Schaik, C., Fox, E., & Sitompul, A. (1996). Manufacture and use of tools in wild Sumatran orangutans. *Naturwissenschaften, 83*, 186–188.

van Schaik, C. P., Isler, K., & Burkart, J. M. (2012). Explaining brain size variation: From social to cultural brain. *Trends in Cognitive Sciences, 16*, 277–284.

Visalberghi, E., & Fragaszy, D. (1990). Do monkeys ape? In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes* (pp. 247–273). Cambridge, UK: Cambridge University Press.

Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P., & Fragaszy, D. (2009). Selection of effective stone tools by wild bearded capuchin monkeys. *Current Biology, 19*, 213–217.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science, 311*, 1301–1303.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology, 5*, e184.

Watts, D., & Mitani, J. (2001). Boundary patrols and intergroup encounters in wild chimpanzees. *Behaviour, 138*, 299–327.

Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behaviour, 86*, 213–221.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences, 11*, 233–273.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B, 364*, 2417–2428.

Wilson, M., Hauser, M., & Wrangham, R. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour, 61*, 1203–1216.

The Evolution of Language

4

Philip Lieberman

Humans are the only living species that possesses language. It is difficult to conceive of any aspect of human behavior in which language does not play a role. Social interaction, mitigating or promoting violence, forming inclusive groups, sharing information, selecting mates, transmitting technology, ethical values all involve communication through the medium of language. And speech, as noted below, has properties that enhance the transmission rate of information. Other means can serve to communicate information, but language is the modality by which human culture is transmitted and information is shared. Language also is a modality of thought that may be thought of as a cultural tool. Current studies point to the neural bases of language and cognition sharing common structures and circuits.

The Darwinian Framework

Charles Darwin knew that the pace of evolution is not even—both gradual changes and abrupt transitions occur. Natural selection, the principle most often associated with Darwin, generally results in small gradual changes. In Darwin's words,

> …any variation, however slight and from whatever cause proceeding, if it be in any degree profitable to an individual of any species, in its infinitely complex relations to other organic beings and to external nature, will tend to the preservation of that individual and will generally be inherited by its offspring. (Darwin 1859, p. 61)

The genetic basis of variation was unknown in Darwin's epoch, but he was acutely aware of the variations that marked the individuals that comprised a species, whether the species in question was pigeons or people. The concept of a species was to Darwin elastic, since his focus was on the "transmutation" of species. Thus, distinction between varieties and species thus necessarily was hazy. For example, when discussing the different varieties of domesticated pigeons, Darwin observed that

> if shown to an ornithologist, and he were told that they were wild birds, would certainly, I think, be ranked by him as well-defined species. (1859, p. 22)

Natural selection, blind to all consequences except the survival of progeny, was the engine that could gradually create a "new" species through the aggregation of small variations. However, Darwin also proposed another mechanism to account for abrupt transitions, such as that from aquatic to terrestrial life. Gradual adaptations by means of natural selection to aquatic life that would produce fish that were more successful at having surviving progeny could not account for the appearance of animals who lived on land and had to breathe air. Darwin's solution,

P. Lieberman (✉)
George Hazard Crooker University,
141 Elton Street, Providence, RI 02906-5433, USA
e-mail: Philip.Lieberman@gmail.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_4,

based on his observation of living species – lungfish, transitional species that could survive in dried up river beds – was the

> …fact that an organ originally constructed for one purpose…may be converted into one for a wholly different purpose…. (p. 190)

Communication and Cognition

It has often been debated whether language serves primarily as a vehicle of thought or communication. The findings of the studies that will be briefly reviewed here point to language serving both functions. Moreover, although human cognitive capabilities are in a sense unique, the neural substrate implicated in human language and cognition also regulates aspects of behavior that we share to various degrees with other species. The biological bases of human language and cognition reflect aspects of anatomy and neural mechanisms that can be traced back millions of years that initially served other ends, but took on new functions and subsequently were adapted through the agency of natural selection to enhance communication and cognition.

It, moreover, is difficult to conceive of any aspect of human behavior in which language does not play a role. Social interaction, mitigating violence, promoting violence, forming inclusive groups, denigrating other groups, sharing information, selecting mates—all have been enhanced by the evolution of human linguistic and cognitive capabilities. That being the case, it is unlikely that any single factor, such as social interaction, provided the selective advantage that drove the evolution of human language and cognition.

Initial Stages

The biological endowment that humans bring to bear on the acquisition of language is innate, an aspect of human genetic endowment. However, children learn to speak, learn the meaning of words, and learn the syntax of their language. This point of view contracts with the view expressed by Noam Chomsky and many other linguists. Chomsky's central claim for decades (e.g., Chomsky 1972) has been that humans possess innate neural mechanisms that are species specific and specific to language alone and that essentially instantiate the details of syntax of all language, past, present, and future. For many years Chomsky posited a "universal grammar" (UG) that instantiated hundreds of principles and parameters that were activated when a child came into contact with a language. In *The Science of Language* (Chomsky 2013), the innate "faculty of language" entails humans possessing only one innate mental capacity "merge" that accounts for humans possessing language as well as cognitive capabilities such as numeracy and arithmetic skills. Merge is defined as the capacity for "…taking two things and putting them together or taking one thing and taking a piece of it and sticking it at the edge" (Chomsky 2013, p. 16). Merge accounts for what Chomsky takes to be the defining feature of human language—being able to construct and comprehend complex sentences that include clauses or conjunctions.

However, Chomsky's views on the evolution of human language unfortunately fall outside the domain of biology in light of his views on natural selection. In Chomsky's words,

> It is perfectly safe to attribute this development [of innate language structures] to "natural selection," so long as we realize that there is no substance to this assertion, that it amounts to no more than a belief that there is some natural explanation for these phenomena. (Chomsky 1972, p. 97)

Discussions of the absence of evidence for natural selection pervade Chomsky (2013), *The Science of Language*. Summing up its view on the evolution of human language,

> you [Chomsky] emphasize in the case of human language—certainly the most distinctive and central mental faculty, one that no other creature has …there is no evidence that a long-term selectional story will work. There are reasons to believe that language was introduced at a single stroke with the introduction of Merge, perhaps some fifty of sixty thousand years ago. (Chomsky 2013, p. 103)

The date for the sudden appearance of human language on different pages of Chomsky (2013) shifts between 50,000 and 100,000 years in the past, but these dates are all implausible in light of evidence from the fossil record, archaeology, and genetics, which points to modern humans and earlier extinct hominins such as the Neanderthals possessing language at least 500,000 years ago, with precursors dating back to the common ancestor of present-day apes and humans six million years ago.

Moreover, "merge" is not a species-specific human attribute. The neural basis for "merge" clearly is present to some degree in many, perhaps all, living species. If the linguistic terminology is stripped away, merge reduces to the associative learning, the process by which animals learn to associate two or more things, acts, and events—and for humans to associate concepts. Ivan Pavlov, in the early years of the twentieth century, showed that dogs can connect the sound of a bell and food. Subsequent studies show that species ranging from mollusks to apes can merge. The mollusks studied by Carew et al. (1981) were very slow at learning to associate electric shocks and a smell pleasant to mollusks, but they did so after several hundred trials.

The mollusk study demonstrated that the neural correlate of associative learning was, as Hebb (1949) had suggested, synaptic modification. Synapses transmit information between "neurons," the basic computing elements of brains. Synapses also code information acquired through associative learning as their "weights." The degree to which they transmit information is modified. In the mollusk study a synapse involved in motor control coded the association between electric shock and smell.

Humans are capable at learning and executing complex acts, concepts, and language that transcend the capabilities of mollusks or any living species. As the studies discussed below demonstrate, these human capabilities involve genetic events in the last 500,000 years that enhanced synaptic connectivity and plasticity. However, the basic mechanism, synaptic plasticity, is present in mollusks and dates back hundreds of millions of years.

Protolanguage

The language capabilities of extinct hominins (species ancestral or related to present-day humans) have been vigorously debated. Chomsky, as noted above, believes that humans possessing language suddenly appeared 100,000 or 50,000 years ago. Mithin (2005) claimed that the Neanderthals, who became extinct about 40,000 years ago, had a language restricted to communicating by humming. Other proposals have claimed that early hominins possessed a "protolanguage" limited to isolated words. These claims can be assessed through comparative studies of ape communication. Present-day apes and humans share a common ancestor. Thus, any aspect of language that apes can master *in an appropriate environment* most likely was present in early hominin languages.

The reference to "environment" is essential. A biological capacity may be present in any species that is not apparent until it is expressed in a particular environment. No one in the year 1805, for example, would have believed that humans could approach each other at closing speeds exceeding 120 miles per hour, separated by three feet or less, and survive while at the same time listening to music or to someone talking, but that occurs routinely on roads throughout the world all day long.

When raised from infancy in a language using environment that makes use of manual gestures or other manual modalities such as pushing buttons on a speech synthesizer, chimpanzees can acquire some aspects of human language. When American Sign Language (ASL) is used by human caretakers to communicate to another and to infant chimpanzees in a fairly normal household, they can acquire active vocabularies of about 150 words and master the inflectional morphology inherent in ASL words (Gardner and Gardner 1969). Formal tests show that chimpanzees also can comprehend distinctions in meaning conveyed by the syntax of spoken simple English sentences (Savage-Rumbaugh et al. 1985). The Gardner chimpanzees also appeared to comprehend spoken English sentences, but no formal tests were administered. It thus is improbable that any hominin species was limited to a "protolanguage" in which they communicated by uttering isolated, single, words or that Neanderthals were limited to humming. However, no nonhuman species can talk. Since apes instead can use sign language and other manual systems to signify words, this lends plausibility to gestures playing a greater role in the early stages of the evolution of language (Hewes 1973).

Talking

Being able to talk entails being able to learn complex, coordinated motor acts involving the lungs, larynx, jaw, lips, soft palate, and the intrinsic and extrinsic muscles of the tongue, and then being able to rapidly and precisely execute these maneuvers. The learning process for human children takes years, contrary to some claims by linguists that language is in place by age 5. Children at age 10 years still are not able to speak at adult rates (Smith 1978). Though communication using manual gestures appears to be simpler since apes can master a limited number of ASL words, formal manual sign languages are an invention dating to the eighteenth century and could not have played a dominant role in the early stages of hominin language claims for the spontaneous invention of complex sign languages (e.g., Nicaraguan Sign Language does not hold up) (Polich 2006).

The obvious advantages of talking over manual gestures are that it isn't necessary to look at the person communicating, it works in the dark, and it frees the talkers' hands so that they can be used to employ tools, carry infants, carry objects, and manipulate devices. In the 1960s it became apparent that human speech has a critical linguistic advantage over all other auditory signals—as we talk we can transmit information at rates that exceed the fusion frequency of the auditory system. Other discrete sound codes, such as Morse code, transmitted at this rate merge into a buzz. Speech achieves this transmission rate by means of the process of "encoding," a term coined by Alvin Liberman and his colleagues (1967). A brief account of the physiology of speech production and some aspects of speech perception is necessary to understand the complexity of this process and why the evolution of the species-specific human tongue plays a critical role in this process.

The source of energy for speech production is the airflow out of the lungs. This, in itself, necessitates a person's executing a set of complex muscular maneuvers, owing to the evolutionary history of the lungs. Darwin (1859, p. 160) noted that that the lungs of air-breathing animals evolved from the swim bladders of fish. Primitive fish, such as sharks, that lack swim bladders have to constantly move to maintain a given depth in the ocean. Swim bladders enabled more "advanced" fish to hover by storing air extracted from water in elastic swim bladders, which adjusted their body's size so as to displace an equal weight of water at a given depth. This allows such fish to hover while expending less energy than fish that must constantly move. Human lungs retain this elastic property; as they expand, they increase the volume of a person's body. No muscle directly acts on the lung sacks to expand them. During quiet inspiration the diaphragm, intercostal, and abdominal muscles expand the space in which the elastic lungs are placed, and the lungs, in turn, expand, storing energy in a manner similar to that of rubber balloons. During expiration, the elastic recoil of the lung sacs, in a manner analogous to that of distended balloons, expels air. At the start of an expiration, where the lungs are at their maximum expansion, the air pressure within the lungs (alveolar air pressure) is at a maximum and then falls as the lungs deflate. The durations of inspiration and expiration during quiet breathing are almost equal.

The alveolar air pressure during expiration impinges on the vocal cords of the larynx. If the vocal cords of the larynx are tensioned and positioned so as to phonate (rapidly opening and closing), a series of "puffs" of air result. The average rate at which these puffs of air occur, the fundamental frequency of phonation (Fo), is perceived as the "pitch" of a speaker's voice. Fo is determined by the magnitude of alveolar air pressure and the tension placed on the vocal cords of the larynx. Since alveolar air pressure is high at the start of a quiet expiration, F0 would start high and rapidly fall when people talk barring the complex compensatory maneuvers that must be unconsciously carried out when we talk.

The human vocal cords are complex structures whose evolution can be traced back to lungfish. A series of adaptations changed their role from sealing the lungs from the intrusion of water to enhancing phonation in different species (Negus 1949). Animals ranging from frogs to humans communicate information by means of variations in F0. In human speech, variations in F0 patterns often signal emotion, but they also transmit referential, linguistic information in most human languages. The Chinese languages, for example, differentiate words by means of F0 patterns specific to given words—the consonant-vowel sequence [ma] in Northern Chinese specifies four different words, differentiated by four different lexical tones (Tseng 1981).

The pattern of alveolar muscular control during speech is quite different so as to produce an alveolar air pressure contour that is almost level throughout a sentence until its end, unless a speaker wishes to accentuate a syllable (Lieberman 1967). The diaphragm is immobilized and the duration of expiration is keyed to the length of the sentence that the speaker intends to produces. Alveolar air pressure is maintained at an almost uniform level by programing a set of instructions to the intercostal and abdominal muscles that can expand the lungs, so that they "hold back" in a graduated manner, keyed to the length of a sentence, against the force generated by the elastic recoil force of the lungs. The alveolar air pressure resulting from the lungs' elasticity otherwise always would be high at the start of the expiration (often blowing the vocal cords apart, ending phonation) and rapidly falls as lung volume falls. The intercostal and abdominal muscles contain muscle "spindles" that can monitor the force that they produce. The diaphragm, which is immobilized during speech and singing, contains few spindles which may account for its being immobilized when we talk (Bouhuys 1974). During spontaneous speech people usually anticipate the length of the sentence that they will produce and inspire more air before the start of a long sentence (Lieberman and Lieberman 1973). It is difficult to establish when hominins might have acquired the ability to control alveolar air pressure during speech. Attempts have been made to relate the size of bony channels through which nerves enervate the diaphragm to alveolar control, but the diaphragm does not enter into speech production.

In most languages and dialects (there are exceptions such as "valley-girl" English), the fundamental frequency of phonation and amplitude of the speech signal remaining more or less level during a sentence, abruptly falling at its end. In many dialects of English and other languages, Fo remains instead level or rises for yes-no questions (Armstrong and Ward 1926; Pike 1945; Lieberman 1967). Frogs signal different states by means of calls that have different F0s. In primates, independent studies such as Cheyney and Seyforth (1990) show that monkeys signal referential information by means of calls that have different F0 contours. This again points out the implausibility of any stage in early hominin communication that exclusively relied on manual gestures.

Mammalian infants and their mothers maintain contact and direct attention using calls that have varying F0 contours—that holds also for humans (Fernald et al. 1989). The neural circuits that control this aspect of phonation appear to have their roots in therapsids, mammal-like reptiles who lived in the age of dinosaurs. The anterior cingulate gyrus (ACC) of the paleocortex is linked by a neural circuit to the basal ganglia, subcortical structures deep within the brain. The circuits appear to be similar in monkeys and humans (Alexander et al. 1998; Lehericy et al. 2004). The inference that the therapsid brain had a neural circuit linking an ACC to the basal ganglia follows from therapsid fossils having the three middle ear bones that characterize all mammals. In true reptiles these bones form a hinge in the jaw, allowing the jaw to open wide so as to swallow large creatures. In the course of evolution, the former three-bone jaw hinge migrated into the mammalian middle ear, serving as a mechanical amplifier that enhances the ability of a mother's and suckling infant's maintaining contact. Lesion studies show that when the neural circuit to the ACC is disrupted in a mouse mother, she will not pay attention to her pups (Maclean and Newman 1982). General problems in maintaining attention occur when neural circuits linking the ACC to basal neural are degraded by Parkinson disease in humans (Cummings. 1993). Maclean and Newman showed that ACC neural circuits also control the laryngeal "mammalian isolation cry." An example is a human infant cry. In human adults, disrupted basal ganglia circuits of the ACC can result in mutism (Cummings 1993) and aberrant patterns of laryngeal control during speech (Lieberman et al. 1992; Pickett et al. 1998).

Thus, it is most likely that hominins, dating back six million years ago to the common ancestor of humans and present-day apes, had some form of spoken language in which words and, most likely, some aspects of syntax were communicated by modulating F0.

The "Gift" of Tongue

Speech encoding which accounts for the rapid rate at which we can vocally transmit information derives from the inherent limit on how the shape and length of the airway above the larynx can be changed (Fig. 4.1).

The airway above the larynx, the supralaryngeal vocal tract (SVT), determines the sound quality of consonants and vowels in a manner analogous to a pipe organ. In a pipe organ a source of acoustic energy that has a wide band of frequencies is filtered by pipes that allow energy to pass through them in narrow ranges of frequency, producing particular musical notes. Phonation, the acoustic energy generated by the larynx, is the "source" of energy for speech sounds such as the vowels and initial consonants of the words *bit* and *map*. The acoustic energy generated by the larynx occurs at the fundamental frequency of phonation and at its harmonics—integral multiples of Fo. For a given shape and length of the SVT, maximum acoustic energy will

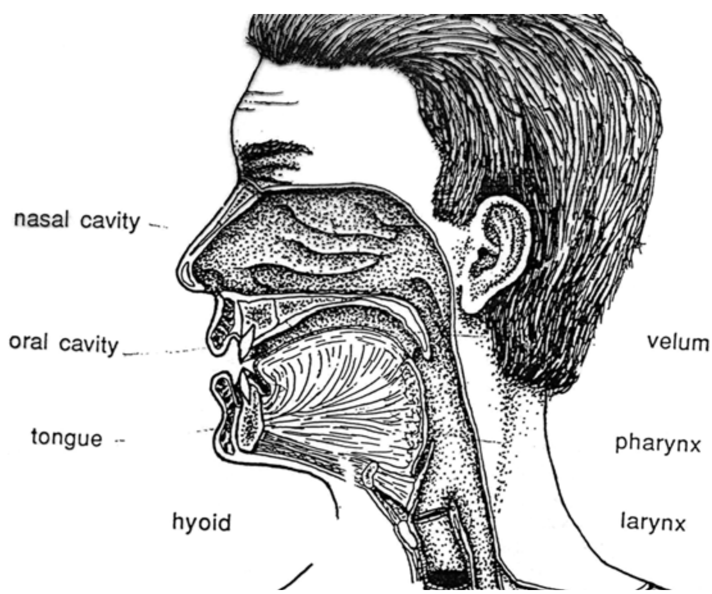


Fig. 4.1 Sketch of the human tongue and supralaryngeal airway

pass through it at a set of "formant frequencies." The vowel of the word "see"—in phonetic notation [i]—could have local energy maxima at 270, 2,300, and 3,000 Hz for an adult speaker whose overall SVT length is 17 cm. The same speaker's [a] vowel (the vowel of the word "ma") would be about 700, 1,100, and 2,400 Hz. Consonants and vowels are distinguished by their formant frequency patterns, as well as their duration, the timing between tongue and lip maneuvers, and laryngeal phonation—detailed discussions can be found in many texts including Fant (1960), Nearey (1978), Lieberman (1984), and Lieberman and Blumstein (1988).

Movable type is often used as an analogy for the discreteness of speech; we hypothetically strung together discrete *phonemes* (sounds approximated by the letters of the alphabet) to form words. The phonemes [t] [a] and [b], for example, could be rearranged to form the words tab, bat, at, and ba. However, discrete phonemes are not present in the acoustic speech signal. In the 1960s it was thought that talking machines could work by simply isolating phonemes from tape recordings of carefully enunciated speech and then rearranging and stringing them together. If phonemes really were beads on a string, there should have been a segment of tape that contained the sound [t], when a person spoke the word *too*, before a segment of tape that contained the sound [u] (the vowel of the word *too*). However, the talking-machine project was a failure; when the segment of the recording tape that was supposed to correspond to the "phoneme" [t] excised from the word *too* was isolated and linked to the hypothetical discrete phoneme of the vowel [i] excised from the word *tea*, the result was incomprehensible.

The dynamics of speech production showed that a talking-machine project based on isolating and linking phonemes never could work. The positions assumed by the lips, tongue, jaw, and larynx for the phoneme /t/ are affected by those necessary to produce the vowel [u] in a different manner than the vowel [i], yielding a different, overall, "encoded" formant frequency patterns for the words *too* and *tea*. The articulatory gestures that generate speech are always encoded unless a person intones a fixed vowel. For example, when producing a word such as bit, the positions of the tongue, jaw, lips, and larynx which determine the shape of the SVT and the resulting formant frequencies of /b/ must move to the different positions necessary for /I/ (the vowel of the word "bit") and then to /t/. As they move, albeit rapidly, there must be a transition between each SVT shape. It became clear that encoded words are the perceptual units of speech, a consonant-vowel (CV) syllable being the minimal unit, that listeners then perceptually decode, taking account of the constraints of speech production at some internal level (Liberman et al. 1967). The formant frequency patterns of the phonemes that make up words are always melded together.

This fact that the phonemes postulated by linguists don't exist in the speech signal allows humans to transmit information at a rate that exceeds that of any other acoustic signal by means of a complex perceptual process. The minimal units transmitted are encoded words, at a rate below 7 units per second. The perceived speech signal then can be perceptually decoded into sequences of phonemes. In a sense, phonemes are abstract speech motor control and perception instruction sets. Chinese orthography, which codes words, is a better approximation of the speech signal than alphabetic systems. Computerized speech recognition systems use algorithms that match the incoming signal to probable word templates. Dogs also can recognize words; some dogs can learn to recognize hundreds of words with one trial, i.e., immediately on hearing the word in association with a referent—usually an object (Kaminski et al. 2004).

One of the problems encountered in speech recognition systems is how to take into account the effect of speakers of all ages and sizes, who have different SVT lengths. Since the length of the SVT varies from person to person and for the same individual as he or she grows up, the absolute values of the formant frequency pattern vary for the same words. For example, the formant frequencies of an /i/ are 1.5 times higher for a child whose SVT length is 11.3 cm longer than for his father whose SVT is 17 cm long. Humans face the same problem, but the two different formant frequency patterns would both be perceived as examples of an /i/ owing to a speech-specific process of perceptual normalization which humans appear to share with other species—dogs can also do it—probably horses and many other species. Listeners internally estimate the length of a speaker's SVT and can estimate SVT length after hearing a short stretch of speech or "reverse-engineering" a known phrase such as a person saying *hello*, but Nearey (1978) showed that the vowel /i/ (of the word *see*) was an optimal signal for SVT normalization. The data of one of the first studies aimed at developing automatic speech recognition (Peterson and Barney 1952) had pointed to the vowel /i/ being an optimal cue for SVT normalization. Two errors in 10,000 trials occurred for words that contained [i] where listeners had to identify monosyllabic words that differed with respect to the vowel and speaker. The words uttered by 10 different speakers were presented in quasi-random order listeners in this experiment, and listeners had to immediately adjust for different speakers' voices. Hillenbrand et al. (1995) using then state-of-the-art computer analysis techniques reported similar results.

Swallowing Versus Speaking

Unfortunately, the human tongue has created a species-specific hazard; compared to other animals, we are more likely to choke on solid food lodged in the larynx. Charles Darwin pointed out the problem:

> The strange fact that every particle of food and drink which we swallow has to pass over the orifice of the trachea, with some risk of falling into the lungs …. (1859, p. 191)

Victor Negus's comparative anatomy showed that the adult human larynx (and opening to the trachea) was carried down into the pharynx because it "is closely apposed to the tongue" (Negus 1949, pp. 25–26). Choking on food, owing to the low human larynx, remains the fourth leading cause of accidental death in the United States (http://www.nsc.org/library/report_injury_usa.htm). Negus thought that the human tongue in some manner facilitated speech communication. That supposition is supported by computer-modeling studies that calculate the range of formant frequencies that can be produced by both nonhuman and human SVTs (e.g., Lieberman et al. 1969, 1972; Lieberman and Crelin 1971; Carre et al. 1995; De Boer 2010).

The initial 1969 study calculated the formant frequency patterns of the vowels that the SVT of a rhesus macaque can produce. The tongue was positioned in the computer model of the monkey's mouth so as to best approximate the SVT configurations used by adult human speakers to yield the vowels [i], [u], and [a]. These vowels delimit the range of vowels used in human languages (Greenberg 1963). The computed monkey vowels did not include these vowels. Newborn infants have SVTs that are similar to those of nonhuman primates (Negus 1949; Crelin 1969). Similar techniques were used to model the SVTs of chimpanzees and human newborn infants (Lieberman et al. 1972). These findings were replicated by Carre et al. (1995) and De Boer (2010). The morphology of the skull of the La Chapelle-aux-Saints Neanderthal fossil indicated that his SVT also was similar to that of a large human newborn (Crelin 1979). Similar computer modeling confirmed that he had a restricted phonetic range that lacked quantal vowels (Lieberman and Crelin 1971). Cineradiographic data of newborn infant cry (Truby et al. 1965) guided the computer modeling of jaw, tongue, lip, and laryngeal maneuvers in these studies.

At birth in humans, the tongue is largely positioned in the mouth; its shape is flat as is the case for other primates and most mammals (Negus 1949). The proportion of the tongue in the oral "horizontal" (SVTh) part of the infant oral cavity relative to the part of the tongue in the "vertical" pharynx (SVTv), SVTh/SVTv, is 1.5, when the larynx is positioned at its lowest point in the "forceful" cries pictured in Truby et al. (1965, pp. 75–78) which were the basis for the computer models used by Lieberman and Crelin (1971) and Lieberman et al. (1972). It is not until age 6–8 years that the human tongue attains its adult 1:1 SVTh/SVTv proportions and almost circular posterior midsagittal shape. The growth process by which the species-specific human vocal tract is formed is complex and sometimes takes as long as 10 years (Lieberman and McCarthy 1999; Lieberman et al. 2001). The length of the oral cavity is first shortened by differential bone growth that moves the hard palate back on the base of the skull (Lieberman 2011). The shape and position of the tongue then gradually change from the newborn tongue, which is flat and is positioned almost entirely in the oral cavity. The human tongue descends down into the pharynx and achieves its posterior rounded contour, carrying the larynx down with it. In contrast, the nonhuman primate tongue throughout life is long, rectangular, and positioned primarily in the oral cavity.

During fetal development and shortly after birth, the chimpanzee larynx drops slightly owing to an increase in the distance between the larynx and hyoid, whereas the human growth pattern involves the continual descent of the tongue (Nishimura 2005). Tongue shape and SVTh/SVTv proportions in nonhuman primates remain almost constant from birth onwards. Darwin's question concerning why humans have a laryngeal position that enhances the propensity to choke thus appears to be an adaptation that enhances the robustness of speech communication.

Stevens (1972) independently showed that only the species-specific human SVT can produce the ten-to-one midpoint area function discontinuities that are necessary to produce the vowels [i], [u], and [a], which Stevens termed "quantal." Stevens employed both computer modeling and physical models (wooden tubes that could be shifted to change the position of the 10:1 changes in SVT cross-sectional area). The quantal vowels are perceptually salient owing to the convergence of two formant frequencies which yield spectral peaks. Their formant frequency patterns also do not shift when tongue position varies about one cm about the midpoint. Speakers thus can be sloppy and produce the "same" vowel. They also can produce the vowel [i] which facilitates the process of vocal tract normalization. The vowel [u] also has this property to a lesser degree—and it takes a human tongue and SVT to be able to produce [i] and [u].

Neural Mechanisms

The neural basis for vocal tract normalization again appears to reflect the mechanism first proposed by Charles Darwin—an organ taking on a new role. Other species appear to estimate the size of conspecifics and other species using the absolute values of their vocalizations' formant frequencies. All other things being equal, larger animals have longer vocal tracts that produce lower formant frequencies. Fitch and Reby (2001) showed that deer lower their larynges so as to lengthen their SVTs, thus increasing the length of their vocal tract and producing lower formant frequencies. The lowered formant frequencies serve to signal to conspecifics that an animal is larger than he actually is. However, the animals' tongues remain anchored in their mouths. The larynx transiently descends by increasing the distance between the hyoid bone and larynx. This maneuver does not change the shape of the SVT—its cross-sectional area functions as a function of distance. The cineradiographs of other mammals vocalizing in Fitch (2000), contrary to Fitch's claims, show that though transient larynx lowering occurs, the animals cannot produce quantal vowels because their tongues are still positioned in their mouths. The dynamic tongue maneuvers discussed in some detail in Fitch (2010, pp. 315–320) do not increase the phonetic range of animal vocalizations. The "formant-dispersion" metric used by Fitch (2000) to infer vocal tract length from formants works only because the animal vocalizations always produce a vowel sound close to the neutral "schwa" vowel of human speech. Their tongues never generate SVT shapes that deviate from a slightly flared tube.

Speech communication for language would be possible without the ability to produce quantal vowels. Lieberman and Crelin pointed it out in their 1971 study which concluded that though Neanderthals lacked the capacity to produce quantal vowels, they could have produced the range of speech sounds other than quantal vowels and some consonants. Neanderthals undoubtedly talked and had some form of language since their stone-working technology, apparent in the archaeological record, could not have been transmitted in the absence of language. In light of the selective advantage speech's information transfer rate, Neanderthals undoubtedly talked, albeit with higher error rates.

The Neural Bases of Language and Cognition

Studies comparing the vocal repertoires of nonhuman primates with SVT modeling studies consistently show that they are unable to make full use of their phonetic potential. They also could communicate to each other or mimic human speech, albeit with reduced intelligibility since their SVTs would enable them to produce all nonquantal vowels and most consonants. However, they cannot talk. Current studies point to a link between the speech motor control capabilities and the cognitive capacities that distinguish humans from other species deriving from neural circuits linking subcortical and cortical structures of the human brain.

The Broca-Wernicke Theory

The traditional answer to the question of why only humans can talk is that Broca's area of the cortex instantiates the "faculty of language." In 1861 Paul Broca published his study of a stroke victim whose speech was limited to a syllable that sounded like *tan*. Broca's postmortem observations were limited to the cortical surface of the brain. Broca's patient's "tan" brain was preserved in alcohol. An MRI of the preserved brain shows that the cortical area usually thought to be Broca's area (the left inferior gyrus) wasn't damaged—an area anterior to it instead was damaged. Moreover, tan had massive damage to the basal ganglia, other subcortical structures, pathways connecting cortical and subcortical neural structures, and other cortical areas (Dronkers et al. 2007).

Nonetheless, in the decades that followed Broca's publication, "Broca's area" has been taken to be the brain's speech and language "organ." A few neurologists demurred, pointing out the fact that postmortem examinations showed that language and speech were disrupted only when subcortical brain damage was present. Neuroimaging studies using CT scans and MRIs have resolved the issue. Damage limited to the cortex including Broca's area, sparing subcortical brain structures, never results in aphasia. Conversely, aphasia can occur when only subcortical structures are damaged (e.g., Alexander et al. (1987)). The conclusion reached by neurologists specializing in aphasia is that it never occurs, absent subcortical damage (Stuss and Benson 1986).

It has become apparent that circuits that link activity in different parts of the brain regulate complex aspects of behavior in both animals and humans. Converging evidence from the deficits of "experiments in nature," such as the stroke that destroyed parts of Paul Broca's patients, neurodegenerative diseases and other insults to the brain, and neuroimaging techniques that monitor activity in the brains of living subjects, points to a class of circuits linking regions of the cortex with the subcortical basal ganglia playing critical roles in motor control, including speech, language, and a range of "higher" cognitive acts. Damage to the basal ganglia or the pathways to it appears to be the basis for the language deficits that characterize aphasia (Lieberman 2000, 2002, 2006).

Tracer studies of the brains of monkeys and other animals that cannot be employed in human studies mapped out these circuits linking areas of the motor cortex with the subcortical basal ganglia. Other cortical-basal ganglia circuits were noted that connected areas of the prefrontal cortex through the basal ganglia and other subcortical structures to temporal and parietal cortical regions of the brain (e.g., Alexander et al. 1986). Noninvasive diffusion tensor imaging (DTI) in human subjects confirms the presence of similar cortical to basal ganglia circuits (Lehericy et al. 2004). Studies of neurodegenerative diseases first identified some of the cognitive and linguistic operations performed by these neural circuits. These circuits are disrupted in Parkinson disease (PD) owing to the depletion of the neurotransmitter dopamine which degrades basal ganglia operations (Jellinger 1990). The basal ganglia act as a sequencing engine, calling out motor control information stored in motor cortex (Marsden and Obeso 1994). PD patients thus have difficulty linking together and executing the submovements of internally directed motor acts that are necessary to walk, talk, or perform manual motor acts (Harrington and Haaland 1991; Speech Lieberman et al. 1992). As Marsden and Obeso pointed out, the basal ganglia in circuits linked to the prefrontal cortex constitute an engine that links cognitive acts and that can shift from one criterion to another as circumstances dictate. Cognitive inflexibility and difficulties occur in cognitive acts that require planning or selecting criteria (e.g., Lange et al. 1992).

The primary cognitive deficit of PD is being unable to change the direction of a thought process or action (Flowers and Robertson 1985). Brain damage limited to the basal ganglia also can result in similar speech and cognitive deficits. Bilateral basal ganglia lesions in the subject studied by Pickett et al. (1998) produced severe speech motor deficits involving sequencing laryngeal, lingual, and lung maneuvers. The subject had difficulty comprehending distinctions in meaning conveyed by syntax and was unable to change the criteria by which she had to sort cards on the "odd-man-out" test, which Flowers and Robertson (1985) devised to test PD patients' cognitive flexibility.

Functional magnetic resonance imaging (fMRI) which monitors local oxygen depletion levels to track the level of neural activity in particular regions of the brain can show when a particular neural structure is active during a cognitive task. Some of the local operations performed in the structures that form human neural circuits in humans thus can be discerned (e.g., Duncan and Owen 2000; Petrides 2005). Dorsal posterior (upper-back) areas of the motor cortex control fine motor control. Ventrolateral (lower-side) prefrontal cortex is active during tasks that involve actively selecting and retrieving information stored in other regions of the brain. The dorsolateral (upper-side) prefrontal cortex is active while monitoring motor or cognitive events during a task taking into account earlier events in working memory. fMRI studies, such as those by Monchi and his colleagues, are revealing some of the roles played by different parts of the basal ganglia and cortex areas in subjects performing linguistic and cognitive tasks. The Wisconsin Card Sorting Test (WCST) is a standard instrument for measuring cognitive flexibility—being able to form and shift from one cognitive criterion to another. The usual form of the WCST involves subjects sorting cards that each have images that differ with respect to shape or color or number. The fMRI data reported in Monchi et al. (2001) shows the activation of a cortical-striatal loop involving the ventrolateral prefrontal cortex, the caudate nucleus of the basal ganglia, and the thalamus when subjects shift to the WCST criterion. A different cortical-striatal loop that includes the posterior prefrontal cortex and the putamen of the basal ganglia is active when a sorting criterion set shift is executed. The dorsolateral prefrontal cortex was involved whenever subjects made any decision as they sorted cards, apparently monitoring whether their responses were consistent with the chosen criterion. Similar activation patterns occurred when subjects were sorting words instead of images and had to match the words on the basis of semantic similarity, the similarity of the beginning of a syllable or rhyme (Simard et al. 2011). The neural circuits involved thus do not appear to be domain specific to either visual or linguistic criteria. Studies ranging from recording electrical activity in basal ganglia neurons of mice and other animals as they learn tasks (Graybiel 1995; Mirenowicz and Schultz 1996; Jin and Costa 2010) to studies of PD patients (Lange et al. 1992, Monchi et al. 2007) and birds (Brainard and Doupe 2002) also show that the basal ganglia in circuits that include cortical areas play a critical role in associative learning and in planning and executing motor acts during speech and, in birds, birdsongs. The starting point may rest in reptiles which possess the basal ganglia. Leaf and Powell (2011) show that the tropical arboreal lizard *Anolis evermanni* exhibits cognitive flexibility when faced with problems fetching food. The question, as is the case for most aspects of evolutionary biology, is the degree of cognitive flexibility that a species possesses.

Fully Human Language and Cognition

The question then arises—why are humans so much better able to perform complex motor acts such as talking, command language, and exhibit the cognitive flexibility that is a key element of human creativity? When and how did these human capacities evolve?

Comparative studies on the architecture of the frontal regions of the brains of monkeys and humans have been conducted over the course of more than a century (Brodmann 1908–1912). Structural cortical differences do not appear to account for nonhuman primates being unable to talk or perform at human cognitive levels. Petrides (2005) in his review of studies of cortical architecture concludes that "the basic architectonic organization (the distribution of neurons in the layers of the cortex) is the same in humans and monkeys." Nor do humans possess the unique cortical to laryngeal neural circuit proposed by Fitch (2012) that hypothetically enables speech motor activity (Lieberman 2012).

A great deal of attention has been focused on humans having big brains because brains require lots of biological support; thus, they must be doing something useful. Current studies show that human brains have about three times as many neurons as a chimpanzee brain (Herculano-Houzel 2009). The posterior temporal cortex is disproportionately larger than would be expected (Semendeferi et al. 1997, 2002). Temporal cortex is part of the brain's long-term information storage system which may explain its larger size. Working memory also may be enhanced by a proportionately larger temporal cortex. Current views on the neural basis of working memory, which involves keeping information in short-term memory during a cognitive process, again suggest that it involves cortical to basal ganglia circuits that access information by means of neural circuits linking prefrontal cortex to temporal cortex and other structures (Badre and Wagner 2006; Postle 2006). The human prefrontal cortex is active in a wide range of cognitive acts and is linked to other neural structures by the basal ganglia, as well as through cortical to cortical circuits. Prefrontal cortical areas appear to pull memory traces of images, words, and probably other stored information out of information-storing regions of the brain (e.g., Postle 2006; Badre and Wagner 2006; Miller and Wallis 2009). Some studies using MRIs have claimed that humans have a disproportionately larger prefrontal cortex than chimpanzees and other apes. However, as Semendeferi et al. (2002) point out, MRIs inherently cannot show that humans have a disproportionately larger prefrontal cortex. It is impossible to differentiate prefrontal cortical areas from the motor regions of the frontal cortex on an MRI; the total human frontal cortex, which includes prefrontal as well as posterior areas involved in motor control, is not disproportionately larger than an ape's.

In short, human brains are bigger, which gives humans the ability to store more information. Memory traces are stored in the temporal cortex and the motor cortex and in parts of the brain that are also associated with processing visual images. Area V1, which is implicated in visual perception, is activated when people think of an image (Kosslyn et al. 1999). It doesn't seem to be the case that any single factor resulted in hominin brains could have driven this process. Abrupt climate changes, alternating periods of glacial cold and heat, have been invoked as stressors that resulted in hominin brains becoming larger. However, there is no evidence for African glacial cold cycles. A different story has invoked alternating periods of drought and heavy rainfall causing alternating desertlike or lush rain forests in the Rift Valley of Africa drive hominin brain size enlargement. However, archaic hominins most likely would have moved away when the climate became cold or desertlike as was the case for other species. Another reoccurring theory posits group size in primates as the causal agent. However, this theory cannot account for solitary orangutans having similar sized brains and gregarious chimpanzees. Given the common neural structures that constitute the neural circuits known to be implicated in motor control, cognition, and language, as well as emotional regulation, it is improbable that any "one" factor was responsible for increasing hominin brain size. Any aspect of behavior that would have led to more progeny surviving—the true test of success in the Darwinian "struggle for existence"—could have contributed to the increase in hominin brain size evident in the fossil record.

Transcriptional Factors

A new avenue of inquiry has opened up that focuses on transcriptional factors. Transcriptional factors essentially are "master" genes that affect the way that other genes are activated to form bodies and brains. The FOXP2 transcriptional factor came to light when the large, extended KE family in London was studied. Severe speech production, sentence comprehension, and cognitive deficits occurred in the family members who had only one copy of the $FOXP2^{human}$ transcriptional factor, instead of the normal two (Fisher et al. 1998). The genetic code in the double helix of DNA has to be transcribed into single-stranded mRNA, which then is translated into proteins and the structures of living organisms. Transcription factors govern this process. The FOXP2 gene is one of many transcription factors. The mouse form of Foxp2 controls the embryonic development of the lungs, the intestinal system, heart, and other muscles, as well as the spinal column (Shu et al. 2001). The structures where FOXP2 and Foxp2 are expressed similar in the human and mouse brains include the cortical-striatal-cortical circuits involved in motor control and cognition—the thalamus, caudate nucleus, and putamen, the neural structures that are anomalous in afflicted members of the KE family (Lai et al. 2003). FOXP2 also acts on the deep layers of the cortex.

The focus on the role of $FOXP2^{human}$ (the superscript signifies the human version of this gene) follows from its being one of the few genes that differ from their chimpanzee version. Humans and chimpanzees share about 98 % of their genes. The human version, $FOXP2^{human}$, evolved during the 6- or 7-million-year period of evolution that separates humans and chimpanzees. In that relative period, $FOXP2^{human}$ underwent two amino acid substitutions in its DNA sequences. Neanderthals and Denisovans, a hominin species that diverged from Neanderthals, and humans all have this version of FOXP2 (Krause et al. 2007; Meyer et al. 2012). Denisovans, who appear to have lived at different times across Eurasia, from Spain to Siberia, either were ancestral to Neanderthals or split off from Neanderthals about 200,000 years ago (Reich et al. 2010). The hominin family tree is becoming very complex, and that question is open.

An additional mutation on a location in the gene (intron 8) which regulates the expression of FOXP2 in neurons is unique to humans (Maricic et al. 2012). This mutation occurred about 200,000 years ago, the period in which anatomically modern human beings appeared and a selective sweep occurred which resulted in $FOXP2^{human}$ spreading throughout the human population (Enard et al. 2002). Selective sweeps occur when a mutation provides a selective advantage that so enhances the survival of individuals and their offspring that it spreads throughout a population over the course of relatively few generations. For example, genes that conferred adult lactose tolerance, providing an additional food source, spread throughout different human groups that herded goats, sheep, or cows (Tishkoff et al. 2007).

Enhanced language, the medium by which virtually all aspects of human technology and culture are transmitted, along with enhanced cognitive ability would account for the 200,000-year intron 8 selective sweep. The mechanisms by which transcriptional factors act are complex, and understanding them is a work in progress, but it is clear that FOXP2 acts to enhance synaptic plasticity in the cortical-basal ganglia circuits that are implicated in speech and cognition (Lieberman 2013a, b). When $FOXP2^{human}$ was knocked into mouse pups, the significant neural difference was increased synaptic plasticity in basal ganglia neurons, as well as increased dendritic lengths in the basal ganglia, thalamus, and layer VI of the cortex (Enard et al. 2009; Reimers-Kipping et al. 2011). The process by which we learn anything involves modifying synaptic "weights"—the degree to which synapses transmit information to a neuron and code information

(Hebb 1949). Other "highly accelerated regions" (HARs) of the genome that are unique to humans also appear to be implicated in neural development (Konopka et al. 2009; Somel et al. 2013). These genetic studies suggest that neural circuits humans share with other primates were, in effect, "supercharged" through the action of transcriptional genes (Lieberman 2013a, b).

Dating Fully Human Language

The genetic evidence briefly reviewed here, the archaeological record, and fossil evidence provide a time line for the evolution of fully human language. Artifacts, in themselves, cannot serve as markers of their maker's cognitive or linguistic abilities. Computer word processing involves a technological base that is far more complex than sharpening the eighteenth-century quill pens, but it does not signify increased human cognitive or literary capabilities. However, there are periods extending over millions of years in which the Oldowan tools attributed to *Homo habilis* are virtually identical. "Hand axes" attributed to *Homo erectus* are more complex; it is difficult to see how the technique necessary to make them could have been transmitted without some form of language, but they conform to the same pattern over 100,000 years. Neanderthal stone tools are still more complex and Neanderthals survived in a cold difficult environment. Their brains were as large as humans and they undoubtedly talked. But one element that characterizes all human cultures was missing—there is almost no trace of the creative impulse that leads humans to produce art—artifacts that are useless. Nor is there evidence of the behavioral "variability"—the pattern of innovation, change, and multiple solutions to a problem that marks human culture (Shea 2011). In contrast, in Africa where modern humans first appeared, a paint "factory" using a complex process that took days to bind pigments to mediums was operating 100,000 years ago (Hensilwood and d'Errico 2011). Stone tools with inscribed decorative lines that date back 75,000 years were found in the same cave. Shell beads dated back 82,000 years were made in North Africa (Bouzouggar et al. 2007).

The time line for the evolution of the unique human tongue suggests an early African origin for fully human cognitive and linguistic capabilities. As Darwin noted, the low position of the human larynx, which follows from the migration of the human tongue down into the neck, creates a choking hazard. A set of acrobatic maneuvers must be carried out to avoid having food falling into and blocking the larynx. The hyoid bone, which supports the larynx, must be pulled forward and upward to move the larynx out of the food path. The epiglottis must simultaneously be flipped down to cover the larynx. The long necks that humans have compared to chimpanzees have a purpose. Human necks must be long so as to accommodate the portion of the human tongue in the neck (SVTv) that's necessary to produce the quantal vowel [i], plus the larynx attached to the tongue body. If the human neck were shorter, it would be impossible to swallow because the larynx would be positioned below it, in the chest where it would be blocked by the collarbone and it would be impossible to swallow.

Robert McCarthy measured the cervical vertebrae of the necks of 73 specimens of modern humans from populations distributed around the globe, as well as a sample of Neanderthals and early modern humans. The long mouths of Neanderthals and earlier extinct hominin species preclude their having the human tongue proportions noted above that can produce quantal vowels. Mouth lengths can be determined with certainty from boney landmarks on fossil skulls (Lieberman and McCarthy in press). If the oral cavity is long, the part of the tongue in the neck must match its length to produce quantal vowels. Neck lengths can be determined from surviving cervical vertebrae.

Tongues that could have produced quantal vowels become evident in Upper Paleolithic fossil hominins unearthed in Europe who lived about 40,000 years ago. Since natural selection can act only on overt behavior, we can be certain that the hominins who had such tongues also had brains that enabled them to talk, else there would have been only the negative consequence of an increased propensity for choking to death that comes from having a human tongue. Thus, we can infer the presence of the brain mechanisms that regulate the voluntary, rapid, complex, internally guided motor acts that underlie human speech. Brains and body coevolved to make human speech possible. But we can infer more than the ability to talk because the cortical-basal ganglia circuits that allow humans to learn and carry out the complex speech motor acts also are part of the human cognitive system.

The stone tools and art of the European Upper Paleolithic have been interpreted as evidence for a "cultural revolution" (Klein 1999), which perhaps led Chomsky (2013) to claim that language suddenly came into being then. However, there would have been no selective advantage for the retention of the peculiar human tongue whose only useful attribute is enhancing the robustness of speech, unless speech was already in place, probably at human levels 200,000 years ago in Africa. Hominin tongues may have initially moved down into the neck as skulls restructured, shortening the mouth and palate (Lieberman 2012), but the neck length of the Skhul V fossil, which has a short mouth (Lieberman and McCarthy 2007, in press), was not long enough to support a fully human tongue, pointing to an increase in neck length leading to the evolution of human speech capabilities. Relative neck length also increases in children between birth and age 6–8 years, as their tongues restructure and move down into the neck (Mahajan and Bharucha 1994).

Conclusion

Some of the biological bases of human language have a long evolutionary history that can be traced to the ancestors of living nonhuman species. Some of the phonetic elements of human language also are used by other species for different ends, such as signaling their size. Chimpanzees have the capacity to acquire and use words and simple syntax. But human capabilities are qualitatively superior.

The neural bases that govern complex aspects of behavior are circuits that link activity in different parts of the brain. Circuits linking regions of the cortex with the basal ganglia regulate internally specified motor acts including speech and a range of cognitive functions, including the ability to change the direction of an act or thought process—cognitive flexibility. Synaptic connectivity and malleability in the basal ganglia have been enhanced by a series of mutations on the FOXP2 transcriptional factor. Selective sweeps point to these mutations playing an important role in the Darwinian "struggle for existence"—the survival of offspring. Other genes unique to humans also appear to have enhanced human cognitive/communicative ability. There is no basis for the sudden appearance of human language 50,000–100,000 years ago without the agency of natural selection.

References

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9*, 357–381.

Alexander, M. P., Naeser, M. A., & Palumbo, C. L. (1987). Correlations of subcortical CT lesion sites and aphasia profiles. *Brain, 110*, 961–991.

Armstrong, L. E., & Ward, I. C. (1926). *Handbook of English intonation*. Leipzig/Berlin: Teubner.

Badre, D., & Wagner, A. D. (2006). Computational and neurobiological mechanisms underlying cognitive flexibility. *PNAS, 103*, 7186–7190.

Bouhuys, A. (1974). *Breathing*. New York: Grune and Stratton.

Bouzouggar, A., Barton, N., Vanhaeren, M., D'errico, F., et al. (2007). 82,000-year-old shell beads from North Africa and implications for the origins of modern human behavior. *PNAS, 104*, 9964–9969.

Brainard, M. S., & Doupe, A. J. (2000). Interruption of a basal-ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature, 404*, 762–766.

Broca, P. (1861). *Nouvelle observation d'aphémie produite par une lésion de la moitié postérieure des deuxième et troisième circonvolutions frontales* (Bulletin Societé Anatomie, 2nd series, Vol. 6, pp. 398–407).

Brodmann, K. (1908). Beiträge zur histologischen Lokalisation der Grosshirnrinde. VII. Mitteilung: Die cytoarchitektonische Cortexgleiderung der Halbaffen (Lemuriden). *Journal für Psychologie und Neurologie, 10*, 287–334.

Carre, R., Lindblom, B., & MacNeilage, P. (1995). *Acoustic factors in the evolution of the human vocal tract. C. R* (Academie des Sciences Paris, t320(Serie IIb), Vol. t320, pp. 471–476).

Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.

Chomsky, N. (1972). *Language and mind*. San Diego: Harcourt Brace Jovanovich.

Chomsky, N. (2012). *The science of language* (Interviews with James McGilvray). Cambridge: Cambridge University Press.

Crelin, E. S. (1969). *Anatomy of the newborn; An atlas*. Philadelphia: Lea and Febiger.

Cummings, J. L. (1993). Frontal-subcortical circuits and human behavior. *Archives of Neurology, 50*, 873–880.

Darwin, C. (1859/1964). *On the origin of species*. Cambridge, MA: Harvard University Press

De Boer, B. (2010). Modeling vocal anatomy's significant effect on speech. *Journal of Evolutionary Psychology, 8*, 351–366.

Dronkers, N. F., Plaisant, O., Iba-Zizain, M. T., & Cananis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain, 130*, 1432–1441.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *TINS, 10*, 475–483.

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature, 41*, 869–872.

Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., et al. (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell, 137*, 961–971.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*, 477–501.

Finlay, B., & Darlington, R. (1995). Linked regularities in the development and evolution of mammalian brains. *Science, 268*, 1578–1584.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localization of a gene implicated in a severe speech and language disorder. *Nature Genetics, 18*, 168–170.

Fitch, W. T. (2000). Skull dimensions in relation to body size in nonhuman mammals: The causal bases for acoustic allometry. *Zoology, 103*, 40–58.

Fitch, W. T. *The evolution of language*. New York: Cambridge University Press.

Flowers, K. A., & Robertson, C. (1985). The effects of Parkinson's disease on the ability to maintain a mental set. *Journal of Neurology, Neurosurgery, and Psychiatry, 48*, 517–529.

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science, 165*, 664–672.

Graybiel, A. M. (1995). Building action repertoires: Memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology, 5*, 733–741.

Greenberg, J. (1963). *Universals of language*. Cambridge: MIT Press.

Harrington, D. L., & Haaland, K. Y. (1991). Sequencing in Parkinson's disease: Abnormalities programming and controlling movement. *Brain, 114*, 99–115.

Henshilwood, C. S., d'Errico, F., van Niekerk, K. L., Coquinot, Y., Jacobs, Z., Lauritzen, S. E., Menu, M., García-Moreno, R., et al. (2011). A 100,000 year old ochre processing workshop at Blombos Cave, South Africa. *Science, 334*, 219–221.

Herculano-Houzel, S. (2009). The human brain in numbers: A linearly scaled-up primate brain. *Frontiers in Human Neuroscience, 3*, 1–11.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology, 14*, 5–24.

Hillenbrand, J. L., Getty, A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099–3111.

Jellinger, K. (1990). New developments in the pathology of Parkinson's disease. In M. B. Streifler, A. D. Korezyn, J. Melamed, & M. B. H. Youdim (Eds.), *Parkinson's disease: Anatomy, pathology and therapy* (Advances in neurology, Vol. 53, pp. 1–15). New York: Raven.

Jin, X., & Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature, 466*, 457–462.

Klein, R. G. (1999). *The human career* (2nd ed.). Chicago: Chicago University.

Konopka, G., Bomar, J., Winden, K., Coppola, G., Jonsson, Z. O., Gao, F., Peng, S., & Preuss, T. M. (2009). CNS development genes by FOXP2. *Nature, 462*, 213–217.

Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., et al. (2007). The derived FOXP2 variant of modern humans was shared with Neandertals. *Current Biology, 17*, 1908–1912.

Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E., & Copp, A. J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain, 126*, 2455–2462.

Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A., & Paul, G. M. (1992). L-dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology, 107*, 394–404.

Leaf, M., & Powell, B. J. (2011). Behavioral flexibility and problem-solving in a tropical lizard. *Biology Letters, 8*, 28–30.

Lehericy, S., et al. (2004). Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Annals of Neurology, 55*, 522–527.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431–461.

Lieberman, P. (1967). *Intonation, perception and language*. Cambridge, MA: MIT Press.

Lieberman, D. E. (2011). *The evolution of the human head*. Cambridge, MA: Harvard University Press.
Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *Journal of Phonetics, 40*, 608–622.
Lieberman, P. (2013a). *The unpredictable species: What makes humans unique*. Princeton: Princeton University Press.
Lieberman, P. (2013b). Synapses, language and being human. *Science, 342*, 944–945.
Lieberman, P., & Crelin, E. S. (1971). On the speech of Neanderthal man. *Linguistic Inquiry, 2*, 203–222.
Lieberman, M. R., & Lieberman, P. (1973). Olson's "projective verse" and the use of breath control as a structural element. *Language and Style, 5*, 287–298.
Lieberman, D. E., & McCarthy, R. C. (1999). The ontogeny of cranial base angulation in humans and chimpanzees and its implications for reconstructing pharyngeal dimensions. *Journal of Human Evolution, 36*, 487–517.
Lieberman, P., & McCarthy, R. M. (2007). Tracking the evolution of language and speech. *Expedition, 49*, 15–20.
Lieberman, P., & McCarthy, R. M. (in press). The evolution of speech and language. In W. Henke & I. Tattersall (Eds.), *Handbook of paleoanthropology* (2nd ed.).
Lieberman, P., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science, 164*, 1185–1187.
Lieberman, P., Crelin, E. S., & Klatt, D. H. (1972). Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the chimpanzee. *American Anthropologist, 74*, 287–307.
Lieberman, P., Kako, E. T., Friedman, J., Tajchman, G., Feldman, L. S., & Jiminez, E. B. (1992). Speech production, syntax comprehension, and cognitive deficits in Parkinson's disease. *Brain and Language, 43*, 169–189.
Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., & Palmer, J. B. (2001). Ontogeny of postnatal hyoid and laryngeal descent: Implications for deglutition and vocalization. *Archives of Oral Biology, 46*, 117–128.
MacLean, P. D., & Newman, J. D. (1988). Role of midline frontolimbic cortex in the production of the isolation call of squirrel monkeys. *Brain Research, 450*, 111–123.
Maricic, T., Gunther, V., Georgiev, O., Gehre, S., et al. (2012). A recent evolutionary change affects a regulatory element in the human FOXP2 gene. *Molecular Biology and Evolution, 30*, 844–852.
McCarthy, R. C., & Strait, D. S. (2005). Morphological constraints on hominin speech production. *Paleoanthropology, 3*, AO2.
Meyer, M., et al. (2012). A high coverage genome sequence from an archaic Denisovan individual. *Science, 338*. doi: 10.1126/science.1224344.
Miller, E. K., & Wallis, J. D. (2009). Executive function and higher order cognition: Definition and neural substrates. In L. Squire (Ed.), *Encyclopedia of neuroscience* (Vol. 4, pp. 99–104). Oxford: Academic.
Mirenowicz, J., & Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature, 379*, 449–451.
Monchi, O., Petrides, M., Petre, K., Worsley, K. J., & Dagher, A. (2001). Wisconsin card sorting revisited; Distinct neural circuits participating in different stages of the test as evidenced by event-related functional magnetic resonance imaging. *The Journal of Neuroscience, 21*, 7739–7741.
Nearey, T. (1978). *Phonetic features for vowels*. Bloomington: Indiana University Linguistics Club.
Negus, V. E. (1949). *The comparative anatomy and physiology of the larynx*. New York: Hafner.
Nishimura, T. (2005). Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. *American Journal of Physical Anthropology, 126*, 193–204.
Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175–184.
Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*, 781–795.
Pickett, E. R., Kuniholm, E., Protopapas, A., Friedman, J., & Lieberman, P. (1998). Selective speech motor, syntax and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: A case study. *Neuropsychologia, 36*, 173–188.
Pike, K. E. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*, 23–38.
Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., Durand, E. Y., et al. (2010). Genetic history of an archaic hominin group from Denisova cave in Siberia. *Nature, 468*, 1053–1060.
Reimers-Kipping, S., Hevers, S., Paabo, S., & Enard, W. (2011). Humanized Foxp2 specifically affects corticobasal ganglia circuits. *Neuroscience, 175*, 75–84.
Savage-Rumbaugh, S., Rumbaugh, D., & McDonald, K. (1985). Language learning in two species of apes. *Neuroscience and Biobehavioral Reviews, 9*, 653–665.
Semendeferi, K., Damasio, H., Frank, R., & Van Hoesen, G. W. (1997). The evolution of the frontal lobes: A volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution, 32*, 375–378.
Semendeferi, K., Lu, A., Schenker, N., & Damasio, H. (2002). Humans and apes share a large frontal cortex. *Nature Neuroscience, 5*, 272–276.
Shea, J. (2011). *Homo sapiens* is as *Homo sapiens* was: Behavioral variability versus "behavioral modernity": In paleolithic archaeology. *Current Anthropology, 52*, 1–35.
Shu, W., Yang, H., Zhang, L., Ju, M. M., & Morrisey, E. E. (2001). Characterization of a new subfamily of winged-

helix/forkhead (Fox) genes that are expressed in the lungs and act as transcriptional repressors. *Journal of Biological Chemistry, 276*, 27488–277497.

Simard, F., Joanette, Y., Petrides, M., Jubault, T., Madjar, C., & Monchi, O. (2011). Fronto-striatal contributions to lexical set-shifting. *Cerebral Cortex, 21*, 1084–1093.

Somel, M., Liu, X., & Khaitovich, P. (2013). Human brain evolution; transcripts, metabolites and their regulators. *Nature Reviews Neuroscience, 14*, 112–127.

Stevens, K. N. (1972). Quantal nature of speech. In E. E. David Jr. & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 51–66). New York: McGraw Hill.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven.

Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., et al. (2007). Convergent adaptation of human lactose persistence in Africa and Europe. *Nature Genetics, 39*, 31–39.

Truby, H. L., Bosma, J. F., & Lind, J. (1965). *Newborn infant cry*. Upsalla: Almquist and Wiksell.

Tseng, C. Y. (1981). *An acoustic study of tones in Mandarin*. PhD dissertation, Brown University.

Tyson, E. (1699). *Orang-outang, sive, Homo sylvestris; or, The anatomy of a pygmie compared with that of a monkey, an ape, and a man*. London: Printed for Thomas Bennet and Daniel Brown.

5 Evolution of the Human Brain: From Matter to Mind

Michel A. Hofman

Introduction

Organisms are faced during their lives with an immense variety of problems, ranging from purely physical ones, such as changes in climate or geomorphic disturbances, to organism-specific problems related to food supply, predation, homeostasis, and reproduction. In order to enhance their chances of survival, organisms have to find adequate solutions for the problems with which they are confronted, for any of them could easily be fatal. Problem solving, in other words, is an essential dynamic survival mechanism, evolved to cope with disturbances in the ecological equilibrium. It can therefore be looked upon as an adaptive capacity enabling organisms to adjust themselves to one another and to their physical environment (see, e.g., Hodos and Campbell 1990; Macphail and Bolhuis 2001; Roth and Wullimann 2001; Reader et al. 2011; Shettleworth 2012a). The organism's adaptability, however, is but one aspect of fitness. Free-moving organisms, for example, can actively explore their environment and thus generate new selection forces that can modify the structures involved. Mayr (1982, p. 612) even argues that "many if not most acquisitions of new structures in the course of evolution can be ascribed to selection forces by newly acquired behaviors."

This suggests that in highly complex organisms, such as primates, behavior rather than environmental change may be the major driving force for evolution at the organismal level. However, this does not detract from the fact that all organisms, whether they are simple reflex automata or active and complex explorers, are above all concerned with keeping track of their local spatiotemporal environment, as part of their struggle for existence. Since sensory information processing and the ability to model reality (or certain parts of it) are essential components in this process, our idea of problem solving seems to correspond reasonably well to the notion of biological intelligence (Hofman 2003). In fact, with the evolution of sensory systems as adaptations to specialized environments, the capacity to process large amounts of sensory information increased and, with that, the power to create more complex physical realities.

In this chapter, some of the organizational principles and operational modes will be explored that underlie the information-processing capacity of the human brain, and it will be argued that the complexity of the cortical network circuitry is a measure of intelligence.

M.A. Hofman, Ph.D. (✉)
Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands
e-mail: M.Hofman@nin.knaw.nl

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_5,

Evolution of the Cerebral Cortex

If we assume that biological intelligence in higher organisms is the product of processes of complex sensory information processing and mental faculties, responsible for the planning, execution, and evaluation of intelligent behavior, variations among species in intelligence must in principle be observable in the neural substrate. In higher organisms, especially in primates, the complexity of the neural circuitry of the cerebral cortex is considered to be the neural correlate of the brain's coherence and predictive power and, thus, a measure of intelligence.

The evolutionary expansion of the cerebral cortex, indeed, is among the most distinctive morphological features of mammalian brains. Particularly in species with large brains, and most notably in great apes and marine mammals, the brain becomes disproportionately composed of this cortical structure (Welker 1990; Nieuwenhuys 1994a, b; Northcutt and Kaas 1995; Striedter 2004; Hofman and Falk 2012; Fig. 5.1). The volume of cortical gray matter, for example, expressed as a percentage of total brain volume increases from about 25 % for insectivores to 50 % for humans (Frahm et al. 1982; Hofman 1988), whereas the relative size of the entire cerebral cortex (including white matter) goes from 40 % in mice to about 80 % in humans (Hofman 1988, 2012; Azevedo et al. 2009; Herculano-Houzel 2009, 2012).

On the other hand, the relative size of the cerebellum remains constant across phylogenetic groups, occupying about 10–15 % of the total brain mass in different orders (Hofman 1988). Comparative studies among four mammalian orders, including primates, have recently revealed that the absolute neuronal composition in the cortex covaries significantly with that of the cerebellum (Herculano-Houzel et al. 2008; Lent et al. 2012), showing that these two brain structures display coordinated growth during phylogenesis in mammals (for reviews, see Herculano-Houzel 2012; Lent et al. 2012).

Such a coordinated evolution of the cerebral cortex and cerebellum fits well with the recent clinical and experimental evidence suggesting an important role of the cerebellum in cognitive and affective functions, in close connection with cortical associative areas (reviewed by Schmahmann 2010). Although the cerebral cortex is not the only brain structure which was selected for in evolution for greater growth, as a result of growing environmental pressure for more sophisticated cognitive abilities, it has played a key role in the evolution of intelligence.

Scaling of the Primate Cerebral Cortex

During the past decades, considerable progress has been made in explaining the evolution of the cerebral cortex in terms of physical and adaptive principles (see, e.g., Macphail and Bolhuis 2001; Hofman 2003; Lefebvre et al. 2004; Lefebvre 2012; Roth and Dicke 2005, 2012). In addition, a quantitative approach to the comparative morphology of the brain has made it possible to identify and formalize empirical regularities in the diversity of brain design, especially in the geometry of the cortex (e.g., Hofman 1989, 2012; Changizi 2001, 2007; Clark et al. 2001).

Analysis of the cerebral cortex in anthropoid primates, for example, revealed that the volume of the neocortex is highly predictable from absolute brain size (Hofman 1989, 2007; Finlay and Darlington 1995; Zhang and Sejnowski 2000; Finlay et al. 2001; for recent reviews see Hofman and Falk 2012). The volume of the cortical gray matter, containing local networks of neurons that are wired by dendrites and mostly unmyelinated axons, is basically a linear function of brain volume, whereas the mass of long-range axons, forming the underlying white matter volume, increases disproportionately with brain size (Fig. 5.2). As a result, the volume of gray matter expressed as a percentage of total brain volume is about the same for all anthropoid primates.

The relative white matter volume, on the other hand, increases with brain size, from 9 % in pygmy marmosets (*Cebuella pygmaea*) to about 35 % in humans, the highest value in primates (Hofman 1989). The nonlinear nature of this

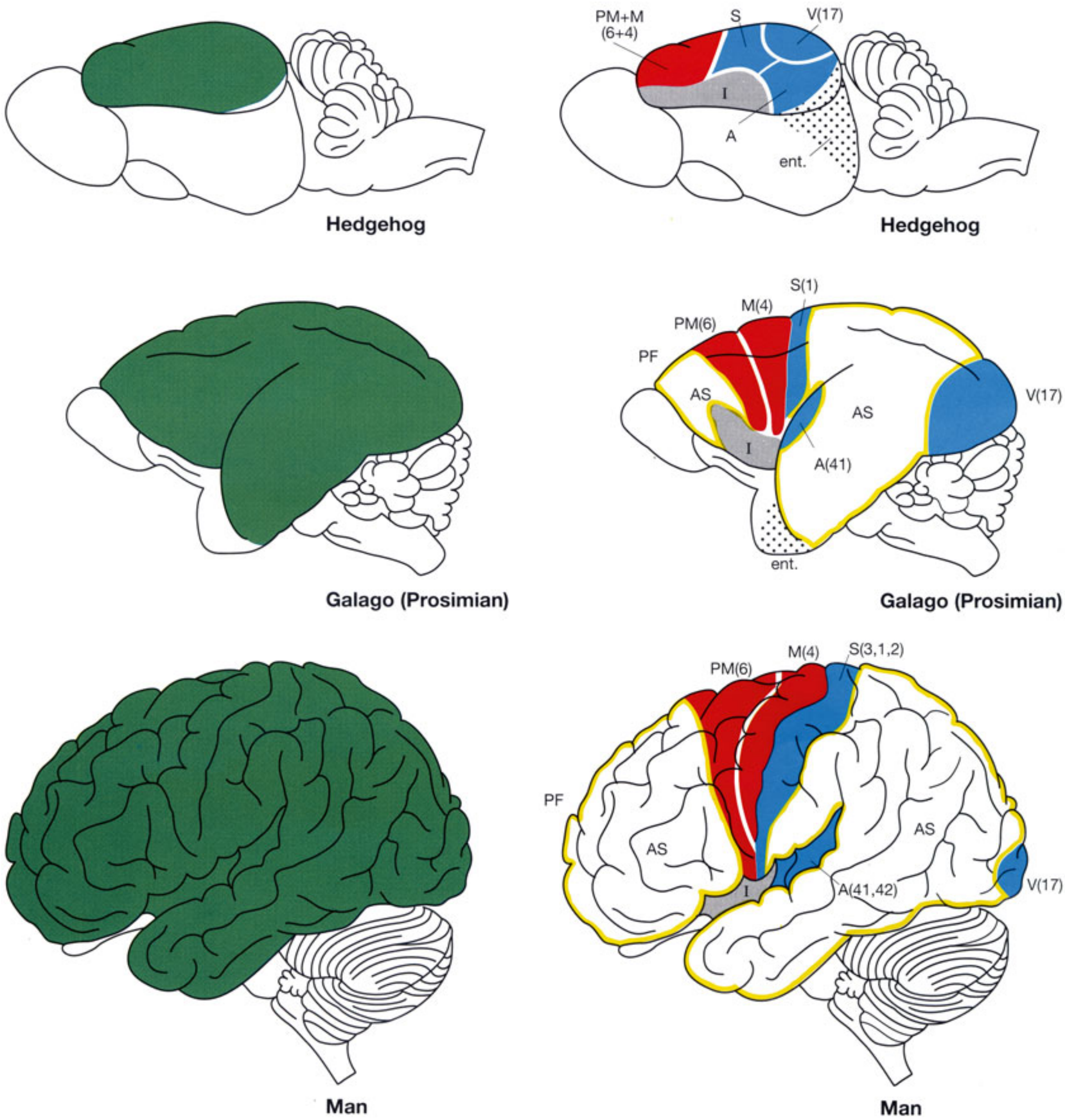


Fig. 5.1 Lateral views of the brains of some mammals to show the evolutionary development of the neocortex (*green*). In the hedgehog, almost the entire neocortex is occupied by sensory (*blue*) and motor (*red*) areas. In the prosimian *Galago*, the sensory cortical areas are separated by an area occupied by association cortex (outlined in *yellow*). A second area of association cortex is found in front of the motor cortex. In man, these anterior and posterior association areas are strongly developed. *A* primary auditory cortex, *AS* association cortex, *Ent* entorhinal cortex, *I* insula, *M* primary motor cortex, *PF* prefrontal cortex, *PM* premotor cortex, *S* primary somatosensory cortex, *V* primary visual cortex (Reproduced with permission from Nieuwenhuys 1994b)

process is further emphasized by plotting the relative volume of white matter as a function of brain size (Fig. 5.3). The high correlation between both variables ensures that the curve, and its confidence limits, can be used for predictive purposes to estimate the volume of white matter relative to brain volume for a hypothetical primate. The model, for example, predicts a white matter volume of about 1,470 cm^3 for an anthropoid primate with a brain volume of 3,000 cm^3 (Hofman 2001b, 2012). In other words, in such a large-brained primate, white matter would

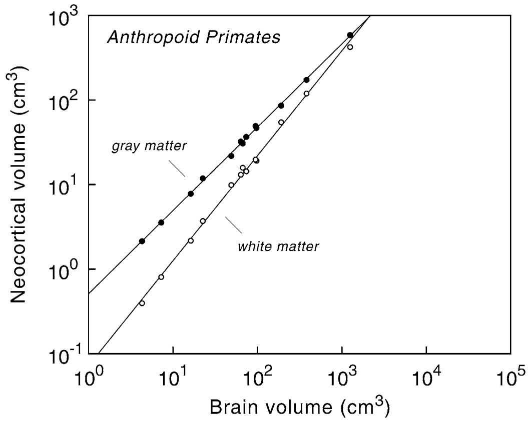


Fig. 5.2 Volumes of cerebral gray and white matter as a function of brain volume in anthropoid primates, including humans. Logarithmic scale. The slopes of the regression lines are 0.985 ± 0.009 (*gray matter*) and 1.241 ± 0.020 (*white matter*). Note the difference in the rate of change between *gray matter* ("neural elements") and *white matter* ("neural connections") as brain size increases (Reproduced with permission from Hofman 2001b)

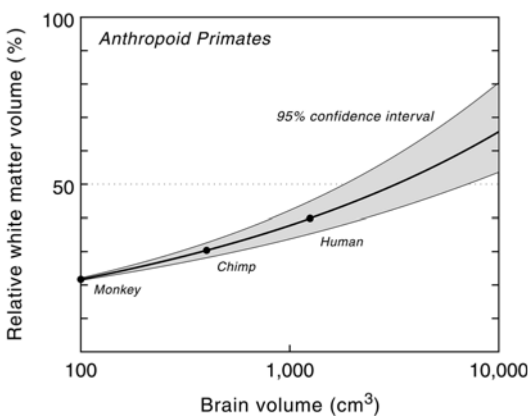


Fig. 5.3 Relative white matter volume as a function of brain volume in anthropoid primates. Semilogarithmic scale. The proportion of white matter increases with brain size, from 22 % in a monkey brain of 100 cm^3 to about 65 % in a hypothetical primate with a brain size of 10,000 cm^3 (Modified with permission from Hofman 2001b)

comprise about half of the entire brain volume, compared to one-third in modern man.

Volumetric measurements of gray and white matter in the neocortex of anthropoid primates have shown that the "universal scaling law" of neocortical gray to white matter applies separately for frontal and non-frontal lobes and that changes in the frontal (but not non-frontal) white matter volume are associated with changes in other parts of the brain, including the basal ganglia, a group of subcortical nuclei functionally linked to executive control (Smaers et al. 2010). These comparative analyses indicate that the evolutionary process of neocorticalization in primates is mainly due to the progressive expansion of the axonal mass that implement global communication, rather than to the increase in the number of cortical neurons and the importance of high neural connectivity in the evolution of brain size in anthropoid primates.

Wen and Chklovskii (2005) have shown that the competing requirements for high connectivity and short conduction delay may lead naturally to the observed architecture of the mammalian neocortex. Obviously, the brain functionally benefits from high synaptic connectivity and short conduction delays. A magnetic resonance imaging study, furthermore, focusing specifically on the prefrontal cortex, has shown that the volume of the white matter underlying prefrontal areas is disproportionately larger in humans than in other primates (Schoenemann et al. 2005). It suggests that the connectional elaboration of the prefrontal cortex, which mediates such important behavioral domains as planning, aspects of language, attention, and social and temporal information processing, has played a key role in human brain evolution.

Design Principles of Neural Organization

Evolutionary changes in the cerebral cortex have occurred mainly parallel to the cortical surface (tangentially) and have been sharply constrained in the vertical (radial) dimension, which makes it especially well suited for the elaboration of multiple projections and mapping systems. A mosaic of functionally specialized areas has indeed been found in the mammalian cortex, some of the functions being remarkably diverse (Kaas 1993, 2008, 2012; Krubitzer 1995, 2007; Schoenemann 2006). At the lower processing levels of the cortex, these maps bear a fairly simple topographical relationship to the world, but in higher areas, precise topography is sacrificed for the mapping of more abstract functions. Here, selected aspects of the sensory input are combined in ways that are likely to be relevant to the animal.

Using modern anatomical tracing methods, physiological recordings, and mapping studies, it has been established that each sensory modality is mapped several times in different areas, with about a dozen representations of the visual world and half a dozen each of auditory inputs and somatosensory sensations. In fact, the maps differ in the attributes of the stimulus represented, in how the field is emphasized, and in the types of computations performed. Clearly, the specifications of all these representations mean that functional maps can no longer be considered simply as hard-wired neural networks. They are much more flexible than previously thought and are continually modified by feedback and lateral interactions. These dynamic changes in maps, which seem likely to result from local interactions and modulations in the cortical circuits, provide the plasticity necessary for adaptive behavior and learning. Although species vary in the number of cortical areas they possess and in the patterns of connections within and between areas, the structural organization of the primate neocortex is remarkably similar.

The tremendous increase in the cortical surface without a comparable increase in its thickness during mammalian evolution has been explained in the context of the radial-unit hypothesis of cortical development (for reviews, see Rakic 2007, 2009). According to this model, neocortical expansion is the result of changes in proliferation kinetics that increase the number of radial columnar units without changing the number of neurons within each unit significantly. Therefore, the evolutionary expansion of the neocortex in primates is mainly the result of an increase in the number of radial columns.

The widespread occurrence of these neocortical columns, furthermore, qualifies them to be considered as fundamental building blocks in neural evolution (for reviews see Mountcastle 1997; Buxhoeveden and Casanova 2002b; Rockland 2010; Buxhoeveden 2012). It has become evident that these cortical circuits integrate at higher levels of information processing, as a result of the hierarchical organization of the brain, thus enabling the system to combine dissimilar views of the world. It implies that if we seek the neural basis of biological intelligence, including mind-like properties and consciousness, we can hardly localize it in a specific region of the brain, but must suppose it to involve all those regions through whose activity an organism is able to construct an adequate model of its external world, perhaps it may even encompass the entire neo- and subcortical network.

It is evident that these neocortical columns are functional and morphological units whose architecture may have been under selective evolutionary pressure in different mammalian lineages in response to encephalization and specializations of cognitive abilities. We are beginning to understand some of the geometric, biophysical, and energy constraints that have governed the evolution of these neural networks (e.g., Felleman and Van Essen 1991; Chklovskii et al. 2002; 2004; Klyachko and Stevens 2003; Laughlin and Sejnowski 2003; Rockland 2010; Casanova et al. 2011). To operate efficiently within these constraints, nature has optimized the structure and function of these processing units with design principles similar to those used in electronic devices and communication networks. In fact, the basic structural uniformity of the cerebral cortex suggests that there are general architectural principles governing its growth and evolutionary development (Cherniak 1995, 2012; Hofman 1996, 2001a, 2007; Rakic 2009; Bullmore and Sporns 2012).

Comparative studies furthermore indicate that variability in subtle subcomponents of the columnar organization in human and nonhuman primates, such as the composition of the interneuron subtypes, are a primary source of interspecific differences in minicolumn morphology among species (Raghanti et al. 2010). Humans deviate from other primates in having a greater width of minicolumns in specific cortical areas, especially in the prefrontal cortex, owing to constituents of the peripheral neuropil space (Buxhoeveden and Casanova 2002a; Semendeferi et al. 2011).

These findings support the idea (Semendeferi et al. 2002; Allen 2009; Teffer and Semendeferi 2012) that human evolution, after the split from the common ancestor with chimpanzees, was accompanied by discrete modifications in local circuitry and interconnectivity of selected parts of the brain. The differences in columnar diameter among primates, however, are only minor compared to the dramatic variation in overall cortex size. Thus, it seems that the main cortical change during evolution has presumably been an increase in the number rather than the size of these neural circuits.

Neural Network Wiring

Although the details of the interpretation of the columnar organization of the neocortex are still controversial (for recent reviews, see Da Costa and Martin 2010; Rockland 2010), it is evident that the potential for brain evolution results not from the unorganized aggregation of neurons but from cooperative association by the self-similar compartmentalization and hierarchical organization of neural circuits and the invention of fractal folding, which reduces the interconnective axonal distances.

Recent network studies, using diffusion tensor imaging (DTI), have demonstrated that the neurons in the neocortex are structurally and functionally highly organized and that this also holds for the wiring of the brain (Van den Heuvel and Sporns 2011; Wedeen et al. 2012). The interconnecting white matter axonal pathways are not a mass of tangled wires, as thought for a long time, but they form a rectilinear three-dimensional grid continuous with the three principal axes of development. The topology of the brain's long-range communication network looks like a 3-D chessboard with a number of highly connected neocortical and subcortical hub regions.

The competing requirements for high connectivity and short conduction delay may lead naturally to the observed architecture of the human neocortex. Obviously, the brain functionally benefits from high synaptic connectivity and short conduction delays. The design of the primate brain is such that it may perform a great number of complex functions with a minimum expenditure of energy and material both in the performance of the functions and in the construction of the system. In general, there will be a number of adequate designs for an object, which, for practical purposes, will all be equivalent.

Recently, we have shown that in species with convoluted brains, the fraction of mass devoted to wiring seems to increase much slower than that needed to maintain a high degree of connectivity between the neural networks (Hofman 2003, 2007). These findings are in line with a model of neuronal connectivity (Deacon 1990; Ringo 1991) which says that as brain size increases, there must be a corresponding fall in the fraction of neurons with which any neuron communicates directly. The reason for this is that if a fixed percentage of interconnections is to be maintained in the face of increased neuron number, then a large fraction of any brain size increase would be spent maintaining such degree of wiring, while the increasing axon length would reduce neural computational speed (Ringo et al. 1994). The human brain, for example, has an estimated interconnectivity of the order of 10^3, based on data about the number of modular units and myelinated nerve fibers (Hofman 2012). This implies that each cortical module is connected to a thousand other modules and that the mean number of processing steps, or synapses, in the path interconnecting these modules is about two.

Herculano-Houzel et al. (2010) have shown that in primates the mass of the white matter scales linearly across species with its number of nonneuronal cells, which is expected to be proportional to the total length of myelinated axons in the white matter. Decreased connectivity in the brain is compatible with previous suggestions that neurons in the cerebral cortex are connected as a small-world network and should slow down the increase in global conduction delay in cortices with larger numbers of neurons (Sporns et al. 2004, 2007; Wang et al. 2008; Fig. 5.4).

Once the brain has grown to a point where the bulk of its mass is in the form of connections, then further increases (as long as the same ratio in interconnectivity is maintained) will be unproductive.

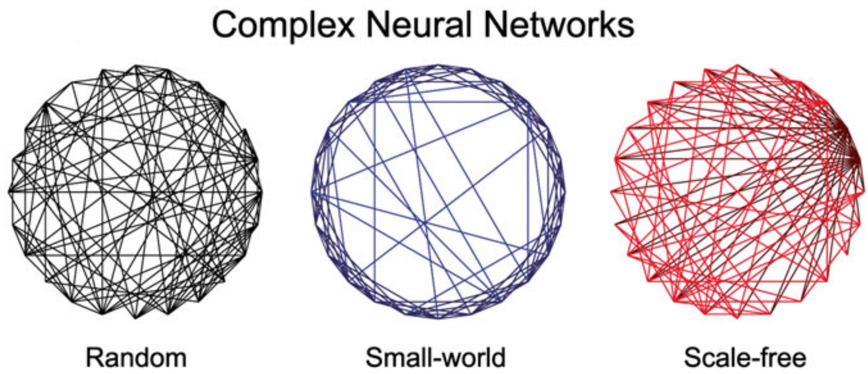


Fig. 5.4 Organizational principles of random, small-world, and scale-free networks. *Structural* cortical networks are neither completely connected with each other nor randomly linked; instead, their connections have small-world attributes with path lengths that are close to those of equivalent random networks but with significantly higher degrees of local clustering. *Functional* cortical networks, on the other hand, exhibit both scale-free attributes with power law degree distributions as well as small-world attributes (Modified with permission from Sporns et al. 2004)

Increases in number of units will be balanced by decreased performance of those units due to the increased conduction time. This implies that large brains may tend to show more specialization in order to maintain processing capacity. Indeed, an increase in the number of distinct cortical areas with increasing brain size has been reported (Welker 1990; Kaas 2000, 2012; Striedter 2004). It may even explain why large-brained species may develop some degree of brain lateralization as a direct consequence of size. If there is evolutionary pressure on certain functions that require a high degree of local processing and sequential control, such as linguistic communication in human brains, these will have a strong tendency to develop in one hemisphere (Ringo et al. 1994; Aboitiz et al. 2003).

Biological Limits to Information Processing

The primate brain, as we have seen, has evolved from a set of underlying structures that constrain its size and the amount of information it can store and process. If the ability of an organism to process information about its environment is a driving force behind evolution, then the more information a system, such as the brain, receives and the faster it can process this information, the more adequately it will be able to respond to environmental challenges and the better will be its chances of survival (Hofman 2003). The limit to any intelligent system therefore lies in its abilities to process and integrate large amounts of sensory information and to compare these signals with as many memory states as possible and all that in a minimum of time. It implies that the functional capacity of a neuronal structure is inherently limited by its neural architecture and signal processing time (see, e.g., Hofman 2001a; Laughlin and Sejnowski 2003; Changizi and Shimojo 2005).

The processing or transfer of information across cortical regions, rather than within regions, in large-brained primates can only be achieved by reducing the length and number of the interconnective axons in order to set limits to the axonal mass (Fig. 5.5). The *number* of interconnective fibers can be reduced, as we have seen, by compartmentalization of neurons into modular circuits in which each module, containing a large number of neurons, is connected to its neural environment by a small number of axons. The *length* of the interconnective fibers can be reduced by folding the cortical surface and thus shortening the radial and tangential distances between brain regions. Local wiring—preferential connectivity between nearby areas of the cortex—is a simple strategy that helps keep cortical connections short.

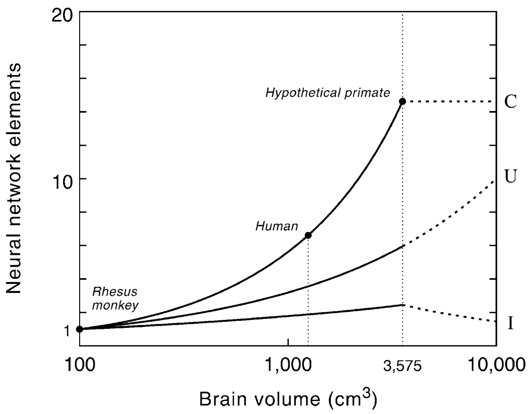


Fig. 5.5 The number of connections (*C*), cortical processing units (*U*), and level of interconnectivity (*I*) in the primate neocortex as a function of brain size. Semilogarithmic scale. Values are normalized to one at a brain volume of 100 cm^3, the size of a monkey brain. Note that the number of myelinated axons increases much faster than the number of cortical processing units (see also Fig. 5.3). The human cerebrum, for example, contains six times more myelinated axons than that of a rhesus monkey, whereas the number of cortical processing units is only three times larger. *Dashed lines* show the potential evolutionary pathway of these neural network elements in primates with very large brains, that is, beyond the hypothetical upper limit of the brain's processing power (see text and Fig. 5.6). Note that a further exponential growth in the number of cortical processing units, without an increase in the number of connections, will lead to a decrease in connectivity between these units and thus to more local wiring (Reprinted with permission from Hofman 2012)

In principle, efficient cortical folding could further reduce connection length, in turn reducing white matter volume and conduction times (Young 1993; Scannell et al. 1995; Chklovskii et al. 2004). Thus, the development of the cortex does seem to coordinate folding with connectivity in a way that could produce smaller and faster brains.

Recently, Wang et al. (2008) have shown that there are functional trade-offs in white matter axonal scaling in mammals. They found that the composition of white matter shifts from compact, slow-conducting, and energetically expensive unmyelinated axons to large, fast-conducting, and energetically inexpensive myelinated axons. The fastest axons have conduction times of 1–5 ms across the neocortex and <1 ms from the eye to the brain, suggesting that in select sets of communicating fibers, large brains reduce transmission delays and metabolic firing costs at the expense of increased volume. Delays and potential imprecision in cross-brain conduction times are especially great in unmyelinated axons, which may transmit information via firing rate rather than precise spike timing. In the neocortex, axon size distributions can account for the scaling of per-volume metabolic rate and suggest a maximum supportable firing rate, averaged across all axons, of 7 ± 2 Hz. Clearly, the white matter architecture must follow a limited energy budget to optimize both volume and conduction time.

Another way to keep the aggregate length of axonal and dendritic wiring low, and with that the conduction time and metabolic costs, is to increase the degree of cortical folding. A major disadvantage of this evolutionary strategy, however, is that an increase in the relative number of gyri can only be achieved by reducing the gyral width. At the limit, the neurons in the gyri would be isolated from the remainder of the nervous system, since there would no longer be any opening for direct contact with the underlying white matter. Prothero and Sundsten (1984) therefore introduced the concept of the gyral "window," which represents the hypothetical plane between a gyrus and the underlying white matter through which nerve fibers running to and from the gyral folds must pass. According to this hypothesis, there would be a brain size where the gyral "window" area has an absolute maximum. A further increase in the size of the brain beyond that point, that is, at 2,800 cm^3, would increase the cortical surface area, but the "window" would decrease, leading to a lower degree of neuronal integration and an increase in response time.

The remarkably high correlation between gray matter, white matter, and brain size in anthropoid primates ensures that the proposed model can be used for predictive purposes to estimate the volume of white matter relative to brain volume for a hypothetical primate (Hofman 2001b). Model studies of the growth of the neocortex at different brain sizes, using a conservative scenario, revealed that with a brain size of about 3,500 cm^3, the total volume of the subcortical areas (i.e., cerebellum,

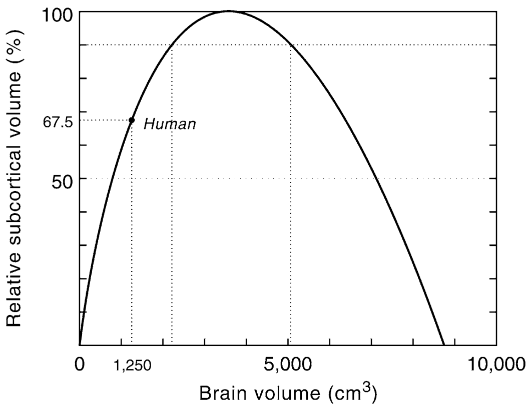


Fig. 5.6 Relative subcortical volume as a function of brain volume. The predicted subcortical volume (i.e., brain volume—predicted neocortex volume) must be zero at zero brain size. Likewise, the subcortical volume will be zero when the brain is exclusively composed of cortical gray and white matter. At a brain size of 3,575 cm^3, the subcortical volume has a maximum (see also Fig. 5.5). The maximum simulated value for the subcortical volume (366 cm^3) is then taken as 100 %. The larger the brain grows beyond this critical size, the less efficient it will become. Assuming constant design, it follows that this model predicts an upper limit to the brain's processing power. Modern humans are at about two-thirds of that maximum (Modified with permission from Hofman 2001b)

brain stem, diencephalon, etc.) reaches a maximum value (Fig. 5.6). Increasing the size of the brain beyond that point, following the same design principle, would lead to a further increase in the size of the neocortex, but to a reduction of the subcortical volume. Consequently, primates with very large brains (e.g., over 5 kg) may have a declining capability for neuronal integration despite their larger number of cortical neurons.

Limits to Human Brain Evolution

A progressive enlargement of the hominid brain started by about 2–2.5 million years ago, probably from a bipedal australopithecine form with a brain size comparable to that of a modern chimpanzee (see, e.g., Falk 2004, 2007, 2012; Robson and Wood 2008; De Sousa and Cunha 2012). The linear scaling law determined for primates allowed Lent et al. (2012) to estimate the number of neurons in the brains of hominins, using brain volumes as inferred from fossil cranial endocasts (Klein 2009). It shows that ancestral primates living between 35 and 20 million years ago—arboricole and quadruped—did not have more than 20 billion neurons in their brains. In the Pliocene period, between 5.3 and 2.5 million years ago, neuronal numbers may have increased to about 40 billion in *Australopithecus*, just above the estimated 30 billion neurons of chimpanzees. These hominins became bipedal and produced the first flaked stone tools. Another increase took place in the early Pleistocene, about 2.5 million years ago, with the appearance of the genus *Homo*. The number of neurons in the brain grew to about 50 billion in *Homo habilis*, reaching about 70 billion in *Homo erectus*, and finally about 90 billion in modern man. With such a large number of neurons, bipedal locomotion consolidated, and hand-finger movements acquired sophisticated abilities, which allowed *Homo* to produce more and more elaborate tools, dominate fire, and improve social interactions. It means that over the past 2–2.5 million years, more than a doubling in the number of neurons has taken place, leading to one of the most complex and efficient structures in the animated universe, the human brain.

In view of the central importance placed on brain evolution in explaining the success of our species, one may wonder whether there are physical limits that constrain its processing power and evolutionary potential. The human brain has evolved from a set of underlying structures that constrain its size and the amount of information it can store and process. In fact, there are a number of related factors that interact to limit brain size, factors that can be divided into two categories: (1) energetic constraints and (2) neural processing constraints (see, e.g., Wang et al. 2008; Herculano-Houzel 2009).

Energetic Limits

The human brain generates about 15 watts (W) in a well-insulated cavity of about 1,500 cm^3. From an engineering point of view, the removal of sufficient heat to prevent thermal overload could be a significant problem. But the brain is actively cooled by blood and not simply by heat conduction from the surface of the head. So the limiting factor is how fast the heat can be removed from the brain by blood flow. It has been suggested by Falk (1990) and others that the evolution of a "cranial radiator" in hominids helped provide additional cooling to delicate and metabolically expensive parts of the brain, such as the cerebral cortex. This vascular cooling mechanism would have served as a "prime releaser" that permitted the brain size to increase dramatically during human evolution. So to increase cooling efficiency in a larger brain, either the blood must be cooler when it first enters the structure or the flow rate must be increased above current levels.

Another factor related to blood flow has to do with the increasing energy requirements of a larger brain, a problem that is exacerbated by the high metabolic cost of this organ. It is unlikely, however, that the rate of blood flow or the increasing volume used by the blood vessels in the brain—in humans about 4 %—constrain its potential size. A bigger brain is metabolically possible because our cardiovascular system could evolve to transport more blood at greater pressure to meet the increased demand. This should not be taken to imply that thermal and metabolic mechanisms play no role at all in setting limits to brain size. Ultimately, energetic considerations will dictate and restrict the size of any neuron-based system, but as theoretical analyses indicate, thermal and metabolic factors alone are unlikely to constrain the potential size of our brain until it has increased to at least ten times its present size (Cochrane et al. 1995).

Neural Processing Limits

The limit to any neural system lies in its ability to process and integrate large amounts of information in a minimum of time, and therefore, its functional capacity is inherently limited by its neural architecture and signal processing time. The scaling model of the geometry of the neocortex, for example, predicts an absolute upper limit to primate brain size (Hofman 2001b; Fig. 5.6). Without a radical change in the macroscopic organization of the brain, however, this hypothetical limit will never be approached, since at that point (ca. 8,750 cm^3), the brain would consist entirely of cortical neurons and their interconnections, leaving no space for any other brain structure.

Cochrane and his colleagues (1995) looked at the different ways in which the brain could evolve to process more information or work more efficiently. They argue that the human brain has (almost) reached the limits of information processing that a neuron-based system allows and that our evolutionary potential is constrained by the delicate balance maintained between conduction speed, pulse width, synaptic processing time, and neuron density. By modeling the information-processing capability per unit time of a human-type brain as a function of interconnectivity and axonal conduction speed, they found that the human brain lies about 20–30 % below the optimal, with the optimal processing ability corresponding to a brain about twice the current volume. Any further enhancement of human brainpower would require a simultaneous improvement of neural organization, signal processing, and thermodynamics. Such a scenario, however, is an unrealistic biological option and must be discarded because of the trade-off that exists between these factors.

Of course, extrapolations based on brain models, such as the ones used in the present study, implicitly assume a continuation of brain developments that are on a par with growth rates in the past. One cannot exclude the possibility of new structures evolving in the brain, or a higher degree of specialization of existing brain areas, but within the limits of the existing "Bauplan," there does not seem to be an incremental improvement path available to the human brain. At a brain size of about 3,500 cm^3, corresponding to a brain volume two to three times that of modern man, the brain seems to reach its maximum processing capacity. The larger the brain grows beyond this critical size, the less efficient it will become, thus limiting any improvement in cognitive power.

Neural Correlates of Consciousness

Consciousness and affective experience may have arisen concurrently in the evolution of the nervous system, as a way to elaborate and extend the potential reach of instinctual urges, while new levels of cortical information processing and cognition promoted the ability of organisms to efficiently pursue goals essential to survival. In fact, affective experience, being an intrinsic brain function, cannot exist independent of consciousness, since in essence it is something that exists as part and parcel of conscious perception (Zeman 2001; Shettleworth 2012a).

In approaching the problem of consciousness, Crick and Koch made the tentative assumption that all the different aspects of consciousness employ a basic mechanism or perhaps a few such mechanisms (Crick and Koch 1990, 1998). In the case of visual consciousness, for example, they have suggested that its biological usefulness in humans is to produce a single but complex interpretation of the visual scene in the light of past experience, either of ourselves or of our ancestors (embodied in our genes), and to make this interpretation directly available, for a sufficient time, to parts of the brain that make a choice among many different but possible plans of action (Crick and Koch 1995). Exactly how this works in detail is unclear.

To be aware of an object or event, Crick and Koch (1995) have argued that the brain has to construct a multilevel, explicit, symbolic interpretation of parts of the visual scene. It means that there are specific groups of neurons at all levels of the visual hierarchy which employ coarse coding to represent some *aspect* of the visual scene. In the case of a particular face, all of these neurons can fire to somewhat face-like objects (Young and Yamane 1992). Notice that while the *information* needed to represent a face is contained in the firing of the ganglion cells in the retina, there is no explicit representation of the face there. A representation of an object or an event will usually consist of representations of many of the relevant aspects of it, and these are likely to be distributed, to some degree, over different parts of the visual system.

The conscious representation of the world is likely to be widely distributed over many areas of the cerebral cortex and possibly over certain subcortical structures as well (Baars 1997). Crick and Koch (1998) postulated that only some types of specific neurons will express the neural correlate(s) of consciousness and that these neurons will probably be fairly close together and will all project roughly to the same place. An alternative hypothesis is that the neural correlate of consciousness is necessarily global (Greenfield 1995). In its most extreme form, this would mean that at one time or another, any neuron in the cortex and associated structures could be part of the neural correlate of consciousness.

The neural correlate of consciousness is defined as the minimal set of neuronal events that gives rise to a specific aspect of a conscious percept (Crick and Koch 2003). The cerebral cortex is probably the most suited part of the brain to look for this neural substrate, as it has very highly and specifically interconnected neuronal networks, many types of excitatory and inhibitory interneurons, and acts by forming transient coalitions of neurons, that is, assemblies of nerve cells, the members of which support one another. The dynamics of coalitions are not simple, as Crick and Koch (1990, 2003) have pointed out. In general, at any moment the winning coalition is somewhat sustained and embodies what an animal is conscious of. On the basis of experimental results in the macaque, Desimone and Duncan (1995) suggest that selective attention biases the competition among competing cell assemblies, but they do not explicitly relate this idea to consciousness.

Coalitions can vary both in size and in character. For example, a coalition produced by visual imagination (with one's eyes closed) may be less widespread than a coalition produced by a vivid and sustained visual input from the environment. These cortical neuronal networks (at least for perception) can be thought of as having nodes. Each node is needed to express one aspect of one percept or another. An aspect cannot become conscious unless there is an essential node for it. For consciousness, there may be other necessary conditions, such as projecting to the frontal cortical areas. Thus, a particular coalition is an active network, consisting of the relevant set of interacting nodes that temporarily sustain itself (Crick and Koch 2003). The smallest useful node may be a cortical column (Mountcastle 1997) or, perhaps, a portion of a cortical column. The feature which that node represents is (broadly) its columnar property. Edelman and Tononi (2000) presented a theory of consciousness, based on the idea of a "dynamic core," which resembles the coalition concept to a large extent. The dynamic core hypothesis, however, rejects the idea that there is a special subset of neurons that alone expresses the neural correlate of consciousness, a view which is also defended in this chapter.

Most of the theories of consciousness have the idea of competing assemblies of neurons in common. Consciousness depends on certain coalitions that rest on the properties of very elaborate neuronal networks. It is suggested that attention consists of mechanisms that bias the competition among coalitions, especially during their formation. Furthermore, the idea that the spatiotemporal dimensions of these nodes represent the neural correlates of mind is most appealing, as it suggests that consciousness, being an integral part of the species' problem-solving capacity, correlates to some extent with the degree of complexity of a nervous system. Therefore, the search for the neural correlates of consciousness should be complemented by a search for its computational correlates (see, e.g., Atkinson et al. 2000; Zeman 2001).

Evolutionary Models of the Mind

Considering biological intelligence as the problem-solving capacity of an organism makes it possible to speak of degrees of intelligence and of its evolution from amoeba to man (Hofman 2003). But what does it mean precisely when one says that species differ in intelligence or that vertebrates are in general more intelligent than invertebrates? It means that there are differences in the abilities of organisms to perceive and interpret the physical world. Biological intelligence can thus be conceived as to reflect the temporal and spatial complexity of the species' niche, without referring, however, to the kinds of situations organisms encounter in everyday life (Reader et al. 2011; Roth and Dicke 2012). It is, in fact, a measure of capacity, independent of the way the capacity is used, and it may be treated as a trait for "anagenetic" rather than "cladistic" analysis (Gould 1976; Jerison 1985). It implies that when distantly related species are comparable in their problem-solving capacity, we should consider the species to be comparable in biological intelligence. Yet the near equality in intelligence may be based upon radically different adaptations. Since neural mechanisms and action patterns evolve in the contexts of the environments in which they are effective, and since species never occupy identical niches, many and various intelligences (in the plural) must have evolved in conjunction with evolving environments (Jerison 1985).

In theory, each ecological niche requires its own degree of biological intelligence. That means that specific neural and sensorimotor adaptations always occur in relation to particular environments. A striking example is the mammalian brain, where the evolutionary changes in the balance of the sensory systems are the result of the adaptive radiation of species into many different ecological niches (Pirlot 1987; Macphail 1993; Macphail and Bolhuis 2001; for reviews, see Roth and Wullimann 2001). These sensory systems, like any other biological feature, could evolve as a result of natural selection, because any subject that forms inadequate representations of outside reality will be doomed by natural selection.

In this view, cognitive systems and emotional phenomena can also be considered to be the result of interactions between genetic aptitude and natural environment, as they have a number of biologically useful functions: one is to keep track of the individual's whereabouts in the world by constructing a schematic model of reality (Popper 1982; Churchland and Churchland 2002; Premack 2007). It is evident that the mind, as an emergent property of sufficiently complex living systems, has its evolutionary history like any other trait that increases adaptation to the environment and that its functions have increased with the evolution from lower to higher organisms (Popper 1982). It might explain the dramatic evolutionary expansion of the human neocortex, being the region where both perception and instruction take place, where the external world is interpreted and modeled, where concepts are formed and hypotheses tested, in short, where the physical world interacts with the mind.

Evolutionary psychology seeks to explain these evolved functional characteristics of the human mind through the lens of an explanatory framework where special adaptive mechanisms are postulated to have been critical for hominid survival and reproductive success (see Panksepp et al. 2002). These "adaptive modules" are theoretical constructs unique to the hominid lineage and should be clearly distinguished from the spatiotemporally defined neural processing units (or modules) of the cerebral cortex discussed in the previous sections (for a review, see Shettleworth 2012b). The existence of a variety of genetically inherited "adaptive modules" is dubious at best when considered simultaneously with our current understanding of mammalian brain organization. Indeed, the organization of the cerebral cortex, which is commonly assumed to be a prime anatomical substrate for unique cognitive functions, exhibits no robust signs of localized anatomical specialization above and beyond specific sensory and motor connections and their polymodal interactions.

Although the adaptation of an organism to its environment is the chief process directing biological evolution, with the evolution of intelligence, organisms became more and more independent of their environments, by modifying the environments according to their needs. This process culminated in the evolution of mankind, which can be understood only as a result of the interaction of two kinds of evolution: the biological and the cultural (Ayala 1986; Herrmann et al. 2007; Premack 2007). Such considerations have led various authors to argue that the human brain can acquire a large variety of epigenetically derived functions via interactions of a limited number of evolutionary conserved affective/motivational systems (situated largely in subcortical areas) with a set of plastic general-purpose learning mechanisms in the cerebral cortex (see Panksepp and Panksepp 2000; Adolphs 2009). It does not mean that there are no special-purpose learning systems in the brain, such as fear learning, but the human cerebral cortex includes much more than a conglomeration of special-purpose learning mechanisms. It contains a neural architecture that can generate flexible features which may be best conceptualized as *rewritable.*

Cultural evolution, however, being the emergent result of the evolution of mind, cannot dispense with biological preconditions; it builds on biological facts and faculties. Though cultural evolution indeed presupposes biological evolution, it is not fully explicable in terms of theories and methods of the latter. In fact, cultural evolution has transgressed organic evolution and shows a certain autonomy (see Hofman 2003). The special status of cultural heredity can be derived from the fact that most cultural innovations are devised precisely in order to meet the environmental challenges or to improve our models of reality, whereas biological evolution has a mindless random character. It is appropriate, therefore, to distinguish adaptations to the environment due to cultural selection from those that take place by the selection of genotypes. Cultural inheritance, furthermore, is an infinitely faster process than genetic inheritance, since it is based on the transmission of information through direct communication and through books, the arts, and the media, which makes that a new scientific discovery, or technical achievement can be transmitted to the whole of mankind in less than one generation (Ayala 1986).

Human Language and Intelligence

It is evident that the role of human language in the transmission of knowledge is extremely important, even so prominent and pervasive that it is hardly possible to estimate human general intellectual capacity independent of linguistic capacity (Macphail 1982; Schoenemann 2012). Its manifestations and, in particular, that of its newly acquired functions—description and argumentation—are the most peculiar phenomena in human problem solving. While animals can communicate by expressing their inner state by means of their behavior and by signaling to conspecifics (e.g., in case of danger), man is the only creature that is able to make true and false statements and to produce valid and invalid arguments.

The progressive accumulation of interactions between environment (both physical and social), “conserved” subcortical systems, and the “general-purpose” cerebral cortex gave rise to a qualitatively different shade of mind—one that could communicate not merely with signs, but in symbolic terms. On the other hand, we have seen that a language system—of the type found in humans—is not essential for consciousness. It is plausible that organisms that do not possess a sophisticated language system are aware of the external world. This is not to say, however, that language does not enrich consciousness or that it does not contribute to our model of reality.

If we assume that part of the basis of human speech is inherited in the DNA and that language is as much a biological as a cultural adaptation, then changes in the brain that permit the advantageous supplement of language acquisition to perception and communication would have had obvious selective advantages throughout the period of hominid evolution (Deacon 1998; Schoenemann 2012). We may conceive human language, therefore, as a superorganic form of adaptation, having evolved not only as a cognitive adaptation contributing to the knowledge of reality of each individual but also as a means of sharing and, even more importantly, influencing states of mind among conspecifics. Indeed, because of language, human beings are not only able to construct individual representations of the external world, but they can also contribute to and learn from *collective* models of reality, that is, the cumulative experience of the whole of mankind. With its cognitive and linguistic skills, *Homo sapiens* tries to know its world and even exerts itself to the utmost to control it.

It is obvious that by virtue of language, human beings tend to have highly organized informational states of mind and, consequently, are excellent problem solvers. But although knowledge of reality may be a necessary condition for survival, it is surely not enough: the degree of intelligence reached by a species does not determine the propensity of its reproductive success. This may be inferred from the indiscriminate elimination of millions of species through the eras, from ammonites to australopithecines. It means that though adaptability increases with the evolution of biological intelligence, environmental catastrophes can always be fatal to a species. But not only external factors can threaten the existence of organisms; *Homo sapiens*, despite its impressive intellectual capacities, might in the end become the victim of its own mind by, paradoxically, creating problems that it is then unable to solve.

Concluding Remarks

All organisms are constantly engaged in solving problems and must therefore have fitting and relevant models of their specific environments in order to enhance their chances of survival. Consequently, the problem-solving capacity of a species is assumed to reflect the temporal and spatial complexity of its ecological niche.

The thesis presented here is that biological intelligence can be considered to be a correlate of the problem-solving capacity of a species, manifesting itself in the complexity of the species' model of reality. With the evolution of sensory systems as adaptations to specialized environments, the capacity to process large amounts of sensory information increased and, with that, the power to create more complex physical realities. The processing of large amounts of information originating from the various sense organs and the construction of complex models of reality require a neural system that selects, integrates, stores, and models—in other words, a system with mind-like properties that enables the organism to make sense of an otherwise chaotic world. But once we allow mind-like properties to come in, such as motivation, emotion, preference, and anticipation, we must allow that it is not only the hostile environment which plays an organizing or designing role in the evolution of biological intelligence but also the active search of an organism for a new ecological niche, a new mode of living.

Since the mind, prehuman and human, takes a most active part in evolution and especially in its own evolution, hominization and the evolution of our linguistic world may have begun as a cultural adaptation to new ecological niches. The process probably started at the time of hominid divergence a few million years ago, as part of the cognitive and manipulative adaptation to what was in essence a more complex physical reality. In other words, some of the seemingly unique higher functions of the human brain, such as language and other neuro-symbolic capacities, were not necessarily due to genetic selection and may have emerged epigenetically through learning and cultural experiences because of the dramatic expansion of the neocortex and its increased tendency to neural plasticity. It seems that the time is finally ripe to begin to building an evolutionary viewpoint of the mind based on comparative concepts that incorporate the intrinsic systems found in all primate brains.

References

Aboitiz, F., Lopez, J., & Mortiel, J. (2003). Long distance communication in the human brain: Timing constraints for inter-hemispheric synchrony and the origin of brain lateralization. *Biological Research, 36*, 89–99.

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology, 60*, 693–716.

Allen, J. S. (2009). *The lives of the brain: Human evolution and the organ of mind*. Cambridge, MA: Belknap.

Atkinson, A. P., Thomas, M. S. C., & Cleeremans, A. (2000). Consciousness: Mapping the theoretical landscape. *Trends in Cognitive Sciences, 4*, 372–382.

Ayala, F. J. (1986). Booknotes. *Biology and Philosophy, 1*, 249–262.

Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. I., Leite, R. E. P., Filho, W. J., Lent, R., & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology, 513*, 532–541.

Baars, B. J. (1997). *In the theater of consciousness*. New York: Oxford University Press.

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience, 13*, 336–349.

Buxhoeveden, D. P. (2012). Minicolumn size and human cortex. *Progress in Brain Research, 195*, 219–235.

Buxhoeveden, D. P., & Casanova, M. F. (2002a). The minicolumn hypothesis in neuroscience. *Brain, 125*, 935–951.

Buxhoeveden, D. P., & Casanova, M. F. (2002b). The minicolumn and the evolution of the brain. *Brain, Behavior and Evolution, 60*, 125–151.

Casanova, M. F., El-Baz, A., & Switala, A. (2011). Laws of conservation as related to brain growth, aging, and evolution: Symmetry of the minicolumn. *Frontiers in Neuroanatomy, 5*, 66. doi:10.3389/fnana.2011.00066.

Changizi, M. A. (2001). Principles underlying mammalian neocortical scaling. *Biological Cybernetics, 84*, 207–215.

Changizi, M. A. (2007). Scaling the brain and its connections. In J. H. Kaas (Ed.), *Evolution of nervous systems* (Vol. 3, pp. 167–180). New York: Academic.

Changizi, M. A., & Shimojo, S. (2005). Parcellation and area-area connectivity as a function of neocortex size. *Brain, Behavior and Evolution, 66*, 88–98.

Cherniak, C. (1995). Neural component placement. *Trends in Neurosciences, 18*, 522–527.

Cherniak, C. (2012). Neural wiring optimization. *Progress in Brain Research, 195*, 361–371.

Chklovskii, D. B., Schikorski, T., & Stevens, C. F. (2002). Wiring optimization in cortical circuits. *Neuron, 34*, 341–347.

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature, 431*, 782–788.

Churchland, P. S., & Churchland, P. M. (2002). Neural worlds and real worlds. *Nature Reviews Neuroscience, 3*, 903–907.

Clark, D. A., Mitra, P. P., & Wang, S. S.-H. (2001). Scalable architecture in mammalian brains. *Nature, 411*, 189–192.

Cochrane, P., Winter, C. S., Hardwick, A. (1995). Biological limits to information processing in the human brain. Retrieved from: http://www.cochrane.org.uk/opinion/archive/articles.phd

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience, 2*, 263–275.

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature, 375*, 121–123.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex, 8*, 97–107.

Crick, F., & Koch, C. (2003). A framework of consciousness. *Nature Neuroscience, 6*, 119–126.

Da Costa, N. M., & Martin, K. A. C. (2010). Whose cortical column would that be? *Frontiers in Neuroanatomy, 4*, 16. doi:10.3389/fnana.2010.00016.

De Sousa, A., & Cunha, E. (2012). Hominins and the emergence of the modern human brain. *Progress in Brain Research, 195*, 293–322.

Deacon, T. W. (1990). Rethinking mammalian brain evolution. *American Zoologist, 30*, 629–705.

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. New York: Norton and Company.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Edelman, G. M., & Tononi, G. A. (2000). *A universe of consciousness*. New York: Basic Books.

Falk, D. (1990). Brain evolution in homo: The "radiator" theory. *Behavioral and Brain Sciences, 13*, 333–381.

Falk, D. (2004). *Braindance: New discoveries about human origins and brain evolution*. Revised and expanded edition. Gainesville: University Press of Florida.

Falk, D. (2007). Evolution of the primate brain. In W. Henke & I. Tattersall (Eds.), *Handbook of palaeoanthropology* (Vol. 2, pp. 1133–1162). New York: Springer.

Falk, D. (2012). Hominin paleoneurology: Where are we now? *Progress in Brain Research, 195*, 253–270.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.

Finlay, B. L., & Darlington, D. B. (1995). Linked regularities in the development and evolution of mammalian brains. *Science, 268*, 1578–1584.

Finlay, B. L., Darlington, D. B., & Nicastro, N. (2001). Developmental structure in brain evolution. *Behavioral Brain Sciences, 24*, 263–278.

Frahm, H. D., Stephan, H., & Stephan, M. (1982). Comparison of brain structure volumes in insectivora and primates. Part I. Neocortex. *Journal für Hirnforschung, 23*, 375–389.

Gould, S. J. (1976). Grades and clades revisited. In R. B. Masterton, W. Hodos, & H. J. Jerison (Eds.), *Evolution, brain and behavior: Persistent problems* (pp. 115–122). Hillsdale: Erlbaum.

Greenfield, S. A. (1995). *Journey to the centers of the mind*. New York: Freeman.

Herculano-Houzel, S. (2009). The human brain in numbers: A linearly scaled-up primate brain. *Frontiers in Human Neuroscience, 3*, 31. doi:10.3389/neuro.09.031.2009.

Herculano-Houzel, S. (2012). Neuronal scaling rules for primate brains: The primate advantage. *Progress in Brain Research, 195*, 325–340.

Herculano-Houzel, S., Collins, C.E., Wong, P., Kaas, J.H. & Lent, R. (2008). The basic nonuniformity of the cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 12593–12598.

Herculano-Houzel, S., Mota, B., Wong, P., & Kaas, J. H. (2010). Connectivity-driven white matter scaling and folding in primate cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 19008–19013.

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360–1366.

Hodos, W., & Campbell, C. B. G. (1990). Evolutionary scales and comparative studies of animal cognition. In R. P. Kesner & D. S. Olton (Eds.), *Neurobiology of comparative cognition* (pp. 1–20). Hillsdale: Erlbaum.

Hofman, M. A. (1988). Size and shape of the cerebral cortex in mammals. Part II. The cortical volume. *Brain, Behavior and Evolution, 32*, 17–26.

Hofman, M. A. (1989). On the evolution and geometry of the brain in mammals. *Progress in Neurobiology, 32*, 137–158.

Hofman, M. A. (1996). Evolution of the human brain: Design without a designer. In N. Elsner & H.-U. Schnizler (Eds.), *Brain and evolution* (pp. 141–170). Stuttgart: Thieme Verlag.

Hofman, M. A. (2001a). Evolution and complexity of the human brain: Some organizing principles. In G. Roth & M. F. Wullimann (Eds.), *Brain evolution and cognition* (pp. 501–521). New York: Wiley.

Hofman, M. A. (2001b). Brain evolution in hominids: Are we at the end of the road. In D. Falk & K. R. Gibson (Eds.), *Evolutionary anatomy of the primate cerebral cortex* (pp. 113–127). Cambridge: Cambridge University Press.

Hofman, M. A. (2003). Of brains and minds: A neurobiological treatise on the nature of intelligence. *Evolution and Cognition, 9*, 178–188.

Hofman, M. A. (2007). Brain evolution and intelligence in primates. In S. Watanabe & M. A. Hofman (Eds.), *Integration of comparative neuroanatomy and cognition* (pp. 33–53). Tokyo: Keio University Press.

Hofman, M. A. (2012). Design principles of the human brain: An evolutionary perspective. *Progress in Brain Research, 195*, 373–390.

Hofman, M. A., & Falk, D. (Eds.). (2012). *Evolution of the brain in primates. From neuron to behavior*. Amsterdam: Elsevier.

Jerison, H. J. (1985). Animal intelligence as encephalization. *Philosophical Transactions of the Royal Society London B, 308*, 21–35.

Kaas, J. H. (1993). Evolution of multiple areas and modules within neocortex. *Perspectives in Developmental Neurobiology, 1*, 101–107.

Kaas, J. H. (2000). Why is brain size so important: Design problems and solutions as neocortex gets bigger or smaller. *Brain and Mind, 1*, 7–23.

Kaas, J. H. (2008). The evolution of the complex sensory and motor systems of the human brain. *Brain Research Bulletin, 75*, 384–390.

Kaas, J. H. (2012). The evolution of neocortex in primates. *Progress in Brain Research, 195*, 91–102.

Klein, R. G. (2009). *The human career* (3rd ed.). Chicago: Chicago University Press.

Klyachko, V. A., & Stevens, C. F. (2003). Connectivity optimization and the positioning of cortical areas. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 7937–7941.

Krubitzer, L. (1995). The organization of neocortex in mammals: Are species differences really so different? *Trends in Neurosciences, 18*, 408–417.

Krubitzer, L. (2007). The magnificent compromise: Cortical field evolution in mammals. *Neuron, 56*, 201–208.

Laughlin, S. B., & Sejnowski, T. J. (2003). Communication in neural networks. *Science, 301*, 1870–1874.
Lefebvre, L. (2012). Primate encephalization. *Progress in Brain Research, 195*, 393–412.
Lefebvre, L., Reader, S. M., & Sol, D. (2004). Brains, innovations and evolution in birds and primates. *Brain, Behavior and Evolution, 63*, 233–246.
Lent, R., Azevedo, F. A. C., Andrade-Moraes, C. H., & Pinto, A. V. O. (2012). How many neurons do you have? Some dogmas of quantitative neuroscience under revision. *European Journal of Neuroscience, 35*, 1–9.
Macphail, E. M. (1982). *Brain and intelligence in vertebrates*. Oxford: Clarendon.
Macphail, E. M. (1993). *The neuroscience of animal intelligence: From the seahare to the seahorse*. New York: Columbia University Press.
Macphail, E. M., & Bolhuis, J. J. (2001). The evolution of intelligence: Adaptive specializations versus general process. *Biological Reviews, 76*, 341–364.
Mayr, E. (1982). *The growth of biological thought. Diversity, evolution and inheritance*. Cambridge, MA: Belknap.
Mountcastle, V. B. (1997). The columnar organization of the brain. *Brain, 120*, 701–722.
Nieuwenhuys, R. (1994a). The neocortex: An overview of its evolutionary development, structural organization and synaptology. *Anatomy and Embryology, 190*, 307–337.
Nieuwenhuys, R. (1994b). The human brain: An introductory survey. *Medica Mundi, 39*, 64–79.
Northcutt, R. G., & Kaas, J. H. (1995). The emergence and evolution of mammalian neocortex. *Trends in Neurosciences, 18*, 373–379.
Panksepp, J., & Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition, 6*, 108–131.
Panksepp, J., Moskal, J. R., Panksepp, J. B., & Kroes, R. A. (2002). Comparative approaches in evolutionary psychology: Molecular neuroscience meets the mind. *Neuroendocrinology Letters, 23*(Special issue, Suppl 4), 105–115.
Pirlot, P. (1987). Contemporary brain morphology in ecological and ethological perspectives. *Journal für Hirnforschung, 28*, 145–211.
Popper, K. R. (1982). The place of mind in nature. In R. Q. Elvee (Ed.), *Mind in nature* (pp. 31–59). San Francisco: Harper and Row.
Premack, D. (2007). Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 13861–13867.
Prothero, J. W., & Sundsten, J. W. (1984). Folding of the cerebral cortex in mammals: A scaling model. *Brain, Behavior and Evolution, 24*, 152–167.
Raghanti, M. A., Spocter, M. A., Butti, C., Hof, P. R., & Sherwood, C. C. (2010). A comparative perspective on minicolumns and inhibitory GABAergic interneurons in the neocortex. *Frontiers in Neuronanatomy, 4*, 3. doi:10.3389/neuro.05.003.2010.
Rakic, P. (2007). The radial edifice of cortical architecture: From neuronal silhouettes to genetic engineering. *Brain Research Reviews, 55*, 204–219.
Rakic, P. (2009). Evolution of the neocortex: A perspective from developmental biology. *Nature Reviews Neuroscience, 10*, 724–735.
Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society London B, 366*, 1017–1027.
Ringo, J. L. (1991). Neuronal interconnection as a function of brain size. *Brain, Behavior and Evolution, 38*, 1–6.
Ringo, J. L., Doty, R. W., Demeter, S., & Simard, P. Y. (1994). Time is of the essence: A conjecture that hemispheric specialization arises from interhemispheric conduction delay. *Cerebral Cortex, 4*, 331–343.
Robson, S. L., & Wood, B. (2008). Hominin life history: Reconstruction and evolution. *Journal of Anatomy, 212*, 394–425.
Rockland, K. S. (2010). Five points on columns. *Frontiers in Neuroanatomy, 4*, 22. doi:10.3389/fnana.2010.00022.
Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in Cognitive Sciences, 9*, 250–257.
Roth, G., & Dicke, U. (2012). Evolution of the brain and intelligence in primates. *Progress in Brain Research, 195*, 413–430.
Roth, G., & Wullimann, M. F. (Eds.). (2001). *Brain evolution and cognition*. New York: Wiley.
Scannell, J. W., Blakemore, C. J., & Young, M. P. (1995). Analysis of connectivity in the cat cerebral cortex. *The Journal of Neuroscience, 15*, 1463–1483.
Schmahmann, J. D. (2010). The role of the cerebellum in cognition and emotion: Personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychology Reviews, 20*, 236–260.
Schoenemann, P. T. (2006). Evolution of the size and functional areas of the human brain. *Annual Review of Anthropology, 35*, 379–406.
Schoenemann, P. T. (2012). Evolution of brain and intelligence. *Progress in Brain Research, 195*, 443–459.
Schoenemann, P. T., Sheehan, M. J., & Glotzer, I. D. (2005). Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nature Neuroscience, 8*, 242–252.
Semendeferi, K., Lu, A., Schenker, N., & Damasio, H. (2002). Humans and great apes share a large frontal cortex. *Nature Neuroscience, 5*, 272–276.
Semendeferi, K., Teffer, K., Buxhoeveden, D. P., Park, M. S., Bludau, S., Amunts, K., Travis, K., & Buckwalter, J. (2011). Spatial organization of neurons in the frontal pole sets humans apart from great apes. *Cerebral Cortex, 21*, 1485–1497.

Shettleworth, S. J. (2012a). *Fundamentals of comparative cognition*. New York: Oxford University Press.

Shettleworth, S. J. (2012b). Modularity, comparative cognition and human uniqueness. *Philosophical Transactions of the Royal Society London B, 367*, 2794–2801.

Smaers, J. B., Schleicher, A., Zilles, K., & Vinicius, L. (2010). Frontal white matter volume in anthropoid primates. *PLoS One, 5*, e9123. doi:10.1371/journal.pone.0009123.

Sporns, O., Chilavo, D. R., Kaiser, M., & Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences, 8*, 418–425.

Sporns, O., Honey, C. J., & Kötter, R. (2007). Identification and classification of hubs in brain networks. *PLoS One, 2*, e1049. doi:10.1371/journal.pone.0001049.

Striedter, G. F. (2004). *Principles of brain evolution*. Sunderland: Sinauer Associates.

Teffer, K., & Semendeferi, K. (2012). Human prefrontal cortex: Evolution, development, and pathology. *Progress in Brain Research, 195*, 191–218.

Van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. *The Journal of Neuroscience, 31*, 15775–15786.

Wang, S. S.-H., Shultz, J. R., Burish, M. J., Harrison, K. H., Hof, P. R., Towns, L. C., Wagers, M. W., & Wyatt, K. D. (2008). Functional trade-offs in white matter axonal scaling. *The Journal of Neuroscience, 28*, 4047–4056.

Wedeen, V. J., Rosene, D. L., Wang, R., Dai, G., Mortazavi, F., Hagmann, P., Kaas, J. H., & Tseng, W. Y. (2012). The geometric structure of the brain fiber pathway. *Science, 335*, 1628–1638.

Welker, W. (1990). Why does cerebral cortex fissure and fold? A review of determinants of gyri and sulci. In E. G. Jones & A. Peters (Eds.), *Cerebral cortex* (Vol. 8B, pp. 23–136). New York: Plenum Press.

Wen, Q., & Chklovskii, D. B. (2005). Segregation of the brain into gray and white matter: A design minimizing conduction delays. *PLoS Computational Biology, 1*, 617–630.

Young, M. P. (1993). The organization of neural systems in the primate cerebral cortex. *Proceedings of the Royal Society of London B, 252*, 13–18.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256*, 1327–1331.

Zeman, A. Z. J. (2001). Consciousness. *Brain, 124*, 1263–1289.

Zhang, K., & Sejnowski, T. J. (2000). A universal scaling law between gray matter and white matter of cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 5621–5626.

Intelligence as a Conceptual Construct: The Philosophy of Plato and Pascal

6

Dana Princiotta and Sam Goldstein

Introduction to Philosophers of Different Times

Philosophers seek truth and meaning in life. Philosophers engage in "the study of ideas about knowledge, truth, the nature and meaning of life" (Merriam-Webster Dictionary 2013). Such an endeavor also entails defining a set of ideas about knowledge, truth, the nature and meaning of life, how to do something, or how to live.

This chapter presents the lives and ideas of whose work sets an integral foundation for our current conceptualizations of intelligence. Plato and Pascal's speculation for the love of wisdom offered hundreds of years ago—the former originating in France and the latter in Greece—remains vibrant today. The two philosophers did not restrict themselves to particular knowledge acquisition. Plato lived to the age of 80 and Pascal 39. During their respective lives, the legacies they bestowed upon the world have impacted the realms of science, religion, morality, mathematics and beyond.

In the Republic, Plato states, "Then you see that this knowledge may be truly called necessary, necessitating as it clearly does the use of the pure intelligence in the attainment of pure truth" (cited in Plato and Jowett 2013, p. 6998). In writing this chapter, we hope to enhance the reader's understanding of Plato and Pascal and their roles in our current conceptualization of intelligence.

Plato

Plato Enters the World

Plato was born circa 428 B.C. in Athens, Greece. Due to the time period of Plato's life, few primary sources exist. Born to parents Ariston and Perictione, it is suggested Plato was named after his grandfather, Aristocles, and acquired the nickname of Plato, based upon the width of his shoulders (Mastin 2008). Ariston passed away and Perictione remarried her uncle. Scholars speculate that Plato had two brothers, Adeimantus and Glaucon; a sister, Potone; and a half brother, Antiphon (Mastin 2008).

D. Princiotta, Ph.D. (✉)
1845 West Orange Grove Road, Tucson, AZ 85704, USA
e-mail: danakprinciotta@gmail.com

S. Goldstein, Ph.D.
University of Utah, 230 South 500 East, Suite 100, Salt Lake City, UT 84102, USA
e-mail: info@samgoldstein.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_6,

Beginnings of a Philosopher

Plato's education commenced in Athens. Plato reportedly learned under the guidance of Socrates beginning at age 20 (Mastin 2008). The two contemplated philosophy in concert for approximately 8 years (408 to 400 B.C.). Plato's studies with Socrates were terminated when Socrates was placed on trial for impiety (e.g., introducing new deities to his society) and the "corruption of youth." Plato would pay homage to his mentor and friend through his Socratic dialogues: Apology, Crito, Euthyphro, and Phaedo.

Following Socrates' execution in 399 B.C., Plato's education ferried him abroad to Italy and Egypt in his studies of geometry, astronomy, and geology as a disciple of Socrates. This travel reportedly took place during 399–387 B.C., approximately. During this period, Plato produced the Apology of Socrates among other written works. Soon thereafter (385 B.C.), Plato opened the doors to the Academy of Athens (credited as the first university) and educated others in philosophy for approximately 40 years. The Academy endured until 539 A.D. The pedagogics of the Academy of Athens emphasized mathematics, politics, philosophy, biology, and astronomy. These areas of study came to be known as Platonism (Mastin 2008). The Byzantine Emperor Justinian I closed the Academy of Athens. The Emperor believed that the Academy was "a threat to the propagation of Christianity" (Mastin 2008).

In his early years, Plato pursued his desire to grasp knowledge, contending that knowledge can either be acquired from a source that already comprehended this knowledge or by acquiring this knowledge independently (Benson 2005). Primarily, the two major outlets in which one could acquire knowledge could be sought through dreams, oracles, and "other divination," or through systematic strategies (Benson 2005). Plato did not believe that knowledge could simply be transferred from one individual to the next, as might a virus. Plato also did not believe that knowledge was acquired in a passive fashion of "consuming" or listening to a lecture (Benson 2005). Plato believed that "in order to acquire genuine knowledge in either of these ways one must engage in a strategy of examination much like the systematic strategy of discovering on one's own" (Benson 2005).

First Efforts of Defining Human Intelligence

In what he called the "first efforts of human intelligence," Plato provides his view as a philosopher (Plato and Jowett 2013):

> They were not, like being or essence, mere vacant abstractions, but admitted of progress and growth, while at the same time they confirmed a higher sentiment of the mind, that there was order in the universe. And so there began to be a real sympathy between the world within and the world without. The numbers and figures which were present to the mind's eye became visible to the eye of sense; the truth of nature was mathematics; the other properties of objects seemed to reappear only in the light of number. Law and morality also found a natural expression in number and figure. Instruments of such power and elasticity could not fail to be a 'most gracious assistance' to the first efforts of human intelligence. (Plato and Jowett 2013, p. 745)

Such efforts are again recognized in Plato's Republic:

> And do you not know, I said, that all mere opinions are bad, and the best of them blind? You would not deny that those who have any true notion without intelligence are only like blind men who feel their way along the road? (Plato and Jowett 2013, p. 6700)

Vehicles in Plato's Philosophy

The gestation of Plato's philosophy relied upon three major factors—reason, will, and appetite (Blaise Pascal 2013). He believed that humans were in a state of conflict between appetites and emotions (Hergenhahn 2009); thus, there is a battle for resources in discussing the needs of the human body and the rational pursuits. "The supreme goal in life, according to Plato, should be to free the soul as much as possible from the adulteration of the flesh" (Hergenhahn 2009, p. 48).

Although Plato believed that intelligence was a trait, he stated:

> By reason of these affections the soul is at first without intelligence, but as time goes on the stream of nutriment abates, and the courses of the soul regain their proper motion, and apprehend the same and the other rightly, and become rational. The soul of him who has education is whole and perfect and escapes the worst disease, but, if a man's education be neglected, he walks lamely through life and returns good for nothing to the world below. (Plato and Jowett 2013, p. 308)

Plato's dialogues allowed him to create and utilize the Socratic method of questioning and answering (Mastin 2008; Plato 2013a, b). One of the major goals in the use of dialogues in his teaching was an attempt to instigate thought within his students. He also aspired to promote the independent thought and independent answering of questions (Mastin 2008). Perhaps the most famous dialogues include the Socratic dialogues, the Symposium, Republic, Timaeus, Critias, and the Sophist, though this list is dependent upon individual interests (Hergenhahn 2009). In this chapter, we attempt to summarize a selection of dialogues as they relate to the construction of intelligence.

Timaeus and Intelligence

Although an accurate chronology of the dialogues does not exist, Timaeus and the Republic share common themes surrounding the construction of intelligence, orderliness, and society. Timaeus is one of Plato's dialogues including Socrates, Timaeus of Locri, Critias, and Hermocrates. The characters or philosophers in this dialogue discuss the physical world and the human beings residing in nature (Taylor 1928).

Plato highlights intelligence as being housed in the soul rather than the body (Carpenter 2008). With the soul separated from the body by his accounts, this lends to suggest that intelligence is viewed as a trait (Carpenter 2008).

> There is further difficulty in explaining this part of the Timaeus—the natural order of thought is inverted. We begin with the most abstract, and proceed from the abstract to the concrete. We are searching into things, which are upon the utmost limit of human intelligence, and then all of a sudden we fall rather heavily to the earth. There are no intermediate steps which lead from one to the other. (Plato and Jowett 2013, p. 807)

Plato explained God's greater plan through Timaeus:

> Timaeus: Why did the Creator make the world? He was good, and therefore not jealous, and being free from jealousy he desired that all things should be like himself. Wherefore he set in order the visible world, which he found in disorder. Now he who is best could only create the fairest; and reflecting that of visible things the intelligent is superior to the unintelligent, he put intelligence in soul and soul in body, and framed the universe to be the best and fairest work in the order of nature, and the world became a living soul through the providence of God. (Plato and Jowett 2013, p. 206)

In other words, the world had become a living entity because the "God" instilled intelligence within the creatures (Taylor 1928). However, this intelligence was to be used for a purpose. "The ancient physical philosophers have been charged by Dr. Whewell and others with wasting their fine intelligences in wrong methods of enquiry; and their progress in moral and political philosophy has been sometimes contrasted with their supposed failure in physical investigations" (Plato and Jowett 2013, p. 774).

Plato's Republic and Intelligence

In Plato's most famous dialogues, Plato's Republic, Plato paints the image of a utopian society. "The Republic is considered one of the single most influential works in the whole of Western Philosophy, although his account of Socrates' trial in the 'Apology' may be the most read" (McClain 1977). Plato's motivation to create a utopian society may have been influenced by the traumatic execution of his mentor, through democratic means, nonetheless (Taylor 1928). Within his society exist three different types of individuals—those prescribing to appetite (e.g., workers, slaves), those prescribing to emotion (e.g., soldiers), and those prescribing to reason (philosopher-kings) (Hergenhahn 2009).

Plato believed that the Republic necessitated leadership by the wise philosophers (Hergenhahn 2009). He explains this by arguing:

> …And so we get four names—two for intellect, and two for opinion—reason or mind, understanding, faith, perception of shadows—which make a proportion—being: becoming:: intellect: opinion—and science: belief:: understanding: perceptions of shadows…And would you have the future rulers of your ideal State intelligent beings, or stupid posts? Certainly not the latter. Then you must train them in dialect, which will teach them to ask and answer questions, and is the coping-stone of the sciences. (Plato and Jowett 2013, p. 1710)

Plato further describes this society:

> The lower portion of the lower or visible sphere will consist of shadows and reflections, and its upper and smaller portion will contain real objects in the world of nature or of art. The sphere of the intelligible will also have two divisions—one of mathematics, in which there is no ascent but all is descent; no inquiring into premises, but only drawing inferences. In this division the mind works with figures and numbers, the image of which are taken not from the shadows, but from the objects, although the truth of them is seen only with the mind's eye; and they are used as hypotheses without being analyzed…. (Plato and Jowett 2013, p. 1453)

He continues by stating, "When the sun shines the eye sees, and in the intellectual world where truth is, there is sight and light. Now that which is the sun of intelligent natures, is the idea of good, the cause of knowledge and truth, yet other and fairer than they are, and standing in the same relation to them in which the sun stands to light" (Plato and Jowett 2013, p. 1419). Within these spheres, Plato delineates four divisions or assignments for knowledge (Plato and Jowett 2013, p. 1435): (1) pure intelligence, (2) active intelligence, (3) faith, and (4) perception of shadows. "And the clearness of the several faculties will be in the same ratio as the truth of objects to which they are related" (Plato and Jowett 2013, p. 1435).

With much effort, Plato attempted to create an order in his society for the philosopher-kings to lead the society; however, Plato admitted to the shortcomings of this Republic:

> You are aware, I replied, that quick intelligence, memory, sagacity, cleverness, and similar qualities, do not often grow together, and that persons who possess them and are at the same time high-spirited and magnanimous are not so constituted by nature as to live orderly and in a peaceful and settled manner; they are driven any way by their impulses, and all solid principle goes out of them. (Plato and Jowett 2013, p. 6646)

The Allegory of the Cave

Within Plato's Republic, we are introduced to his allegory of the cave. In his allegory of the cave, Plato discusses our method of acquiring information, namely, the education of society. In this allegory, Plato paints an image of prisoners confined to a cave in which they can only look forward (Plato and Jowett 2013). Prisoners are able to view shadows of individuals passing behind them through light from a fire (Hergenhahn 2009). Plato depicts the prisoners' understanding of their world as a lower form of understanding from the viewpoint of the Divided of Line previously discussed (Hergenhahn 2009). The shadows provide the current reality for the prisoners.

After providing various scenarios including the release of the prisoners from their confines, Plato concludes that the reality of the prisoners would certainly be shifted by the introduction of the sun and their new reality. Furthermore, "if they could lay hands on the man who was trying to set them free and lead them up, they would kill him" (Trotter 1986, p. 257; Hergenhahn 2009). This is an attempt to create an analogy in which educators are leading individuals into the light we identify as education. Leading the prisoner into the light is seen as leading an individual into the world of reason rather than simple sensory experiences (Hergenhahn 2009). In other words, those Plato describes as ignorant would not wish to be enlightened by reason, by his account.

Pythagorean Concepts

It is commonly stated that the second half of Plato's career was heavily influenced by the Pythagoreans (Taylor 1928). With the assistance of the Pythagoreans, Plato argued that experiences of the mind were superior to those experienced by the flesh (Hergenhahn 2009). “Plato borrowed much from the Pythagoreans. It was through the Platonic philosophy that elements of the Dionysiac-Orphic religion became part of the heritage of Western civilization” (Hergenhahn 2009, p. 35).

One of the prominent concepts Plato adopted from the Pythagoreans in the second half of his academic “career” entailed the Theory of Forms. Plato proposes a “dualistic universe consisting of abstract forms and matter” (Hergenhahn 2009, p. 665). This concept states, “Everything in the empirical world is a manifestation of a pure form (idea) that exists in the abstract. Thus chairs, chariots, rocks, cats, dogs, and people are inferior manifestations of pure forms” (Hergenhahn 2009, p. 46). Stated differently, most of what we “experience” is merely a manifestation of thoughts, ideas, or forms of the actual “object” (Hergenhahn 2009, p. 46). Furthermore, “The only true knowledge is that of the forms, a knowledge that can be gained only by reflecting on the innate contents of the soul. Sensory experience interferes with the attainment of knowledge and should be avoided” (Hergenhahn 2009, p. 64). In other words, knowledge of intelligence relies on “intuitive” grasping of the world, not the grasping of pure matter or pure form—“the material world as it seems to us is not the real world, but only a shadow or a poor copy of the real world” (Mastin 2008).

A second important Pythagorean concept, Reminiscence Theory of Knowledge, states that “all knowledge is innate and can be attained only through introspection, which is the searching of one’s inner experiences” (Hergenhahn 2009, p. 48). Plato was a firm believer in reincarnation and reason being a function of the immortal soul (Hergenhahn 2009). Furthermore, Plato believed “All nature is akin, and the soul has learned everything, so that when a man has recalled a single piece of knowledge…there is no reason why he should not find out all the rest, if he keeps a stout heart and does not grow weary of the search, for seeking and learning are in fact nothing, but recollection” (Hamilton and Cairns 1961, p. 364).

Aristotle and the Future of Plato’s Work

Aristotle was one of Plato’s best-known students (Taylor 1928). For Aristotle, it was challenging to follow Plato and his teachings at times.

> Plato’s departure from the simpler realism of his master, as noted by Aristotle, towards that “intelligible world” opposed by him so constantly to the visible world, into which many find it so hard to follow him at all, and in which the “ideas” become veritable persons. To speak, to think, to feel, about abstract ideas as if they were living persons; that is the second stage of Plato’s speculative ascent. (Plato and Jowett 2013, p. 1774)

Others carried on Plato’s teaching. They were known as Neoplatonists (Mastin 2008). As Western civilization continued to develop, Plato’s original writings were lost but resurfaced through the help of Greek Neoplatonists (McClain 1977).

Legacy

Plato died in approximately 348 B.C. in Athens, Greece, in his early 80s (Blaise Pascal 2013). He provided foundations for the future of mathematics, reason, morals, equality, science, and nature. Western society has inherited Plato’s teachings of the physical body and the immortal soul, which has provided support for the prominent world religions including Jewish, Christian, and Islamic studies. Plato’s teachings are imperative to our modern construction of intelligence, making him a prominent figure for educators, historians, and scientists to study.

Pascal

Pascal Enters the World

Blaise Pascal was born on June 19, 1623, in Clermont-Ferrand, France, to Etienne and Antoinette Pascal (O’Connor and Robertson 1996). He was a brother to sisters Jacqueline and Gilberte (Lataste 1911). His father assisted in the courts of Clermont as an advocate and was also known to be a tax collector and mathematician. The family moved to Paris for educational opportunities for the children following Antoinette's passing (Blaise Pascal 2013).

A Child Prodigy Is Recognized

In Paris, Blaise would learn Latin, Greek, and Spanish (Lataste 1911) from his father. Although his father wished to delay Blaise's acquisition of mathematics, Blaise taught himself mathematical principles (Blaise Pascal 2013). He understood that the sum of the angles of a triangle is always 180° at a very young age (O'Connor and Robertson 1996). Blaise's development continued to climb sharply, and his father recognized Blaise as a child prodigy. Contributing to the fields of mathematics and science at an early age following homeschooling in religion, philosophy, math, and physics (O'Connor and Robertson 1996), Blaise was inducted into the Academie Libre, a group focusing on the development of science and mathematics (Blaise Pascal 2013). At the young age of 16, Blaise proposed and defended mathematical theorems. Shortly thereafter, he published his *Essay on Conic Sections* (Lataste 1911).

Major Academic Contributions and Inventions

In his short 39 years, the *Essay on Conic Sections* was merely the beginning of important theorems relevant today. Geometry and probability were two particularly important fields of study and contribution for Pascal (Lataste 1911). The "mystic hexagram" was one of his earlier theorems, developed in adolescence as well (O'Connor and Robertson 1996), otherwise known as Pascal's theorem. His theorems in mathematics would later influence Sir Isaac Newton in his own work relating to general binomial theorem for fractional and negative powers (Blaise Pascal 2013).

Pascal's inventions included an arithmetic machine (i.e., a calculator) by the age of 18, dubbed the Pascaline. He later invented the hydraulic press and syringe (O'Connor and Robertson 1996), the roulette machine (1655), and the cycloid (1659), an instrument able to calculate solid volumes (O'Connor and Robertson 1996, 2007). He was infinitely interested in atmospheric pressure including measurements of atmospheric pressure through estimations of weight (Blaise Pascal 2013).

Pascal's Ailing Health

By 1646, Pascal's ailing health impeded his work. Pascal's physician recommended he rest for a period of time (Lataste 1911). At this point in time, his beliefs and behaviors were altered:

> There are two vectors of experience in Pascal's world, vectors that for inattentive people would have remained unconnected but that for a creative mind might well come together in new and interesting ways. First of all, in this time of his life, Pascal was a man of the world, living up to the expectations of his class living the life of an intelligent gentleman. This meant parties; this meant diversion; this meant gambling. It is unlikely, given his health that he engaged in any dangerous liaisons nor would his conscience have allowed it. So instead, he invented the new mathematics of probability, the calculation of expectations and a mathematics for probing the mysterious heart of Fortune. (Connor 2006, p. 3679)

As stated, these life experiences afforded him opportunities to develop the probability theorem with the help of his colleague Pierre de Fermat (Rogalsky 2006; O'Connor and Robertson 1996). Pascal apprehended that there was indeed a fixed likelihood of a particular outcome when rolling a dice in 1654 (Blaise Pascal 2013). This probability theorem was recorded in Pensées or "Thoughts" and contributed to Pascal's Wager, discussed below.

The Influence of a Near-Death Experience

Pascal's work was channeled in a different manner when he had a frightening near-death experience in the year 1654 (O'Connor and Robertson

1996). This experience fueled two of Pascal's greatest legacies—Lettres Provinciales and Pensées (Lataste 1911). Pascal described himself in the following manner:

> What the gentlemen found after a time was a young man whose intelligence far outstripped their own, who was able to see further and think more clearly, who had a pungent wit, who could be cruel one minute and solicitous the next. They came to know a terribly unhappy young man who wanted to be a saint and yet love life, who berated himself for his worldliness one minute and then laughed at some raucous joke the next. If they had been told that this Pascal fellow would soon have a mystical experience of God and withdraw from the world, they would not have been surprised, for he was halfway there anyway. If they had been told that he would soon after write one of the greatest and nastiest works of the modern age, the Provincial Letter, a book of satire that Voltaire kept by his bedside every night and used as a model, and that these letters would be directed against the all-powerful Jesuits, they would not have been surprised by that, either. (Connor 2006, p. 354)

Pascal's Wager

Previously known as a man of science, Pascal became known as a man of God—though in reality he was both (Lataste 1911). He proposed Pascal's Wager, in that God is rational—"If God does not exist, one will lose nothing by believing in him, while if he does not, one will lose everything by not believing" (Pascal 1660; O'Connor and Robertson 1996). In general, Pascal insisted that mankind must gamble (O'Connor and Robertson 1996). Simply it stated: "If I wager for and God is-infinite gain; if I wager for and God is not-no loss. If I wager against and God is-infinite loss; if I wager against and God is not-neither loss nor gain" (Pascal 1660; Lataste 1911).

Pascal's insistence on wagering extends beyond the religious domain. When discussing religion, he states that in choosing whether God exists or not, "A game is being played at the extremity of this infinite distance where heads or tails will turn up. What will you wager? According to reason, you can do neither the one thing nor the other; according to reason, you can defend neither the one thing nor the other; according to reason, you can defend neither of the propositions" (Pascal 1660; Pascal & Trotter 1910).

This argument is one in which Pascal remains a famous philosopher three and a half centuries later. His words resonate with us today:

> —Yes; but you must wager. It is not optional. You are embarked. Which will you choose then; Let us see. Since you must choose, let us see which interests you least. You have two things to lose, the true and the good; and two things at stake, your reason and your will, your knowledge and your happiness; and your nature has two things to shun, error and misery. Your reason is no more shocked in choosing one rather than the other, since you must of necessity choose…. In truth, there is infinity between the certainty of gain and the certainty of loss. But the uncertainty of the gain is proportioned to the certainty of the stake according to the proportion of the chances of gain and loss. (Pascal 1660; Pascal & Trotter 1910)

This is in accord with other "thoughts on mind and on style" when Pascal states, "People are generally better persuaded by the reasons which they have themselves discovered than by those which have come into the mind of others" (Pascal 1660; Hughes 2011). This is true in Pascal's own life experiences with a near-death experience and religious enlightenment.

Pensées

Pascal reveals his thoughts on intelligence and reason in Pensées. He applied this to science, but in later years of life, to religion (Lataste 1911). As a philosopher, Pascal initiated important discussions regarding intelligence and the acquisition of knowledge. In Pensées (Pascal 1910), Pascal discusses the realm of mathematics and the acquisition of principles, distinguishing the process of acquisition among individuals. Pascal stated:

> The reason, therefore, that some intuitive minds are not mathematical is that they cannot at all turn their attention to the principles of mathematics. But the reason that mathematicians are not intuitive is that they do not see what is before them, and that, accustomed to the exact and plain principles of mathematics, and not reasoning till they have well inspected and arranged their principles, they are lost in matters of intuition where the principles do not allow of such arrangement. (Pascal 1660; Pascal & Trotter 1910, p. 2)

He continues by stating:

> And thus it is rare that mathematicians are intuitive and that men of intuition are mathematicians, because mathematicians wish to treat matters of intuition mathematically and make themselves ridiculous, wishing to begin with definitions and then with axioms, which is not the way to proceed in this kind of reasoning. Not that the mind does not do so, but it does it tacitly, naturally, and without technical rules; for the expression of it is beyond all men, and only a few can feel it. (Pascal 1660; Pascal & Trotter 1910, p. 2)

Furthermore, "Mathematicians who are only mathematicians have exact minds, provided all things are explained to them by means of definitions and axioms; otherwise they are inaccurate and insufferable, for they are only right when the principles are quite clear" (Pascal 1660; Pascal & Trotter 1910, p. 2).

In other words, Hughes (2011) interprets Pascal's position as the difference between intuitive and mathematical minds laying within their perspective—"All mathematicians would then be intuitive if they had clear sight, for they do not reason incorrectly from principles known to them; and intuitive minds would be mathematical if they could turn their eyes to the principles of mathematics to which they are unused" (Pascal 1660; Hughes 2011). In conclusion, "The greater intellect one has, the more originality one finds in men. Ordinary persons find no difference between men" (Pascal & Trotter 1910, p. 3). Pascal further elaborates in perhaps a judgmental tone, "Do you wish people to believe good of you? Don't speak" (Pascal 1660; Hughes 2011).

Intelligence and Religion

On the flip side to the previous examples from Pensées, an example is illustrated in Pensées number 423—"the heart has its reasons of which reason knows nothing" (Pascal 1660; Pascal & Trotter 1910, p. 172; Rogalsky 2006). Similarly, "Those who are accustomed to judge by feeling do not understand the process of reasoning, for they would understand at first sight and are not used to seek for principles. And others, on the contrary, who are accustomed to reason from principles, do not at all understand matters of feelings, seeking principles and being unable to see at a glance" (Pascal 1660; Rogalsky 2006). Ultimately, Pascal melds the two with Pensées 172: "The way of God who disposes all things with gentleness, is to instill religion into our minds with reasoned arguments and into our hearts with grace, but attempting to instill it into hearts and minds with force and threat is to instill not religion but terror" (Pascal 1660; Rogalsky 2006).

Final Thoughts on Pascal's Philosophy

Pascal believed that philosophy was a process, as most philosophers or scientists might also conclude. He argued the following: "Skepticism. Excess, like defect of intellect is accused of madness. Nothing is good but mediocrity" (Pascal 1660; Hughes 2011). Further, "All the principles of philosophers are true: the skeptics, the stoics, the atheists and so on. But their conclusions are false, because the opposite principles are also true" (Pascal 1660; Hughes 2011). He argues that philosophers like Descartes probe science too deeply, describing him as "useless and uncertain" (Pascal 1660; Rogalsky 2006). Pascal challenged the likes of Descartes and the "Cartesian Monad" alike:

> There are different kinds of right understanding; some have right understanding in a certain order of things, and not in others, where they go astray. Some draw conclusions well from a few premises, and this displays an acute judgment. For example, the former easily learn hydrostatics, where the premises are few, but the conclusions are so fine that only the greatest acuteness can reach them. And in spite of that these persons would perhaps not be great mathematicians, because mathematics contain a great number of premises, and there is perhaps a kind of intellect that can search with ease a few premises to the bottom: and cannot in the least penetrate those matters in which there are many premises. Others draw conclusions well where there are many premises. (Pascal 1660; Pascal & Trotter 1910)

Pascal's Legacy

Blaise Pascal struggled with dyspepsia (i.e., digestive disorder) as well as insomnia for half of his life. He died on August 19, 1662, of a malignant stomach tumor in Paris at the age of 39 (Lataste 1911). He died two months after his 39th birthday. As a final gesture, he invited a Parisian family to live in his home and left the estate to this homeless family when he passed.

During his last year of life, he also focused on helping the economically less fortunate travel through Paris by developing and crafting a plan for individuals to travel with horse-drawn carriages for a modest fare (Rogalsky 2006). According to sources, the profits were then donated to charity. At the time, this was the first mass-transit system in Paris (Rogalsky 2006).

Concluding Remarks

Although it is difficult to locate all of Plato's primary sources, most of his 36 dialogues and 13 letters are available through different avenues on the World Wide Web. The very nature of information science can be traced back to Plato and Socrates' philosophy regarding intelligence and the acquisition of knowledge. We might credit Plato for teaching us how to think. Notice the statement "teaching us *how* to think" rather than "teaching us *what* to think."

Plato instituted the first institute of higher education in the Western world, an amenity that many continue to derive educational wealth from. Although political causes attempted to thwart Plato's drive as a philosopher (the first being the execution of his mentor, Socrates, early in life), Plato continued as a philosopher, and the academy prevailed for hundreds of years under the guidance of continuing scholars (Carpenter 2008). The academy continued for centuries following Plato's death, under the guidance of his nephew, Speusippus, with Aristotle on staff as a teacher for 20 of those years (Carpenter 2008).

The hope is as depicted in Plato's Republic, "And so we arrive at the result, that the pleasure of the intelligent part of the soul is the pleasantest of the three, and that he of us in whom this is the ruling principle has the pleasantest life" (Plato and Jowett 2013, p. 7935).

Certainly Plato laid the framework for Pascal, but as we have seen, Pascal was an independent thinker and contributed a wealth of innovative ideas. Pascal influenced many fields of study, including physics, computer science, and geometry. Three hundred years beyond Pascal's death, in the 1970s, a unit of measurement was named after him related to atmospheric pressure, the Pascal (Pa). We continue to benefit from Pascal's logic. He is best remembered as a French Roman Catholic philosopher (Rogalsky 2006). Awards are given in France in Pascal's honor, known as the Blaise Pascal Chairs for international scientists in the Ile-de-France region. A university in his hometown of Clermont-Ferrand is named after Pascal. A programming language for computer scientists has also been named after Pascal. Finally, Pascal's law refers to a principle in hydrostatics that is still utilized (O'Connor and Robertson 1996).

Pensées was considered to be one of Pascal's greatest contributions. Unfortunately, his work was incomplete at his time of death (O'Connor and Robertson 1996). He accomplished much in his 39 years. Pascal has challenged us to better understand and weight the merits of both reason and religion (Rogalsky 2006). He accomplished this though burdened by illness for much of his adult life. He has challenged and motivated three hundred years of philosophers and thinkers (Rogalsky 2006). T. S. Eliot paid homage to Pascal by describing him as "a man of the world among ascetics, and an ascetic among men of the world" (Broome 1965). According to documents, Pascal stated even on his deathbed, "Sickness is the natural state of Christians" (Broome 1965).

The legacies of Plato and Pascal are undeniable. The memory of these men, though difficult to locate primary sources, is alive within the works of philosophers and scientists existing between the time periods of their lives into the twenty-first century. Both men produced important contributions and philosophies to the concepts of intelligence.

We close with words from Pascal. Pascal wrote:

> True eloquence makes light of eloquence, true morality makes light of morality; that is to say, the

morality of the judgment, which has no rules, makes light of the morality of the intellect. For it is to judgment that perception belongs, as science belongs to intellect. Intuition is the part of judgment, mathematics of intellect. To make light of philosophy is to be a true philosopher. (Pascal 1660; Jowett 2013)

References

Benson, H. H. (2005). Plato's rationalistic method. In A. Nelson (Ed.) A companion to rationalism . Oxford, UK: Blackwell Publishing Ltd. doi: 10.1002/9780470996904.ch5.

Blaise Pascal. (2013). *The biography channel website*. Retrieved 22 July 2013, from http://www.biography.com/people/blaise-pascal-9434176

Broome, J. H. (1965). *Pascal*. London: Edward Arnold.

Carpenter, A. D. (2008). Embodying intelligence: Animals and us in Plato's Timaeus. In Dillon, J., & Zovko, J. (Eds.). Berlin: Academie Verlag.

Connor, J. A. (2006). *Pascal's wager: The man who played dice with God*. San Francisco: HarperSanFrancisco.

Hergenhahn, B. R. (2009). *An introduction to the history of psychology* (6th ed.). Wadsworth: Cengage Learning.

Hughes. (2011). Blaise Pascal thoughts (Pensees). Retrieved 16 Sept 2013 from http://sqapo.com/pascal.htm

Lataste, J. (1911). Blaise Pascal. In *The Catholic encyclopedia*. New York: Robert Appleton Company. Retrieved 26 Sept 2013 from New Advent: http://www.newadvent.org/cathen/11511a.htm

Mastin, L. (2008). Socratic—by historical period—the basics of philosophy. Web. 01 Oct 2013.

McClain, E. G. (1977). *The Pythagorean Plato: Prelude to the song itself*. Stony Brook/Boulder: N. Hays.

Merriam-Webster. (2013). Merriam-Webster.com. Retrieved 1 Oct 2013, from http://www.merriam-webster.com/dictionary/philosophy

O'Connor, J. J., & Robertson, E. F. (1996). Blaise Pascal. Retrieved 22 July 2013, from http://www-history.mcs.st-andrews.ac.uk/Biographies/Pascal.html

O'Connor, J. J., & Robertson, E. F. (2007). *A history of calculus*. University of St Andrews. Consultado el, 07–08.

Pascal, B. (1910). *Pascal's Pensées* (trans: Trotter, W. F.).

Pascal, B. (1660). *Pensées* (trans: Trotter, W. F.). New York: Collier & Son, 1910.

Plato. (1986). (W. F. Trotter, Trans.). Roslyn, NY: Black.

Plato. In Hamilton, E., & Cairns, H. (1961). *The collected dialogues of Plato, including the letters*. New York: Pantheon Books.

Plato. (2013a). *The biography channel website*. Retrieved 30 Sept 2013, from http://www.biography.com/people/plato-9442588

Plato. (2013b). *Essential dialogues of Plato* (trans: Jowett, B.). Digireads.com Publishing.

Plato, & Jowett, B. (2013). *Plato's the republic*. Plato (trans: Jowett, B.). ISBN 9781775420514: Amazon.com: Books. Web. 01 Oct 2013.

Rogalsky, T. (2006). Blaise Pascal: Mathematician, mystic, disciple. Journal of the Association of Christians in Mathematical Sciences. Retrieved 02 Oct 2013 from http://www.acmsonline.org/journal2006.htm.

Taylor, A. E. (1928). *A commentary on Plato's Timaeus*. Oxford: Clarendon Press.

7 The Life and Evolution of Early Intelligence Theorists: Darwin, Galton, and Charcot

Jordan Rigby

Charles Robert Darwin

Charles Robert Darwin was born on February 12, 1809, to a world with evolutionary theory still in its infancy. Thinkers before him, such as Plato and Aristotle, did not particularly believe in evolution (Hergenhahn 2009). This did not prevent theorists from building on what was available. In fact, the year Darwin was born Jean Lamarck published a book suggesting an inheritance of acquired traits relating to the passing down of muscular characteristics from adult to offspring. Contributions such as this identified what was missing from early theories, and a transformation took place. Another contributor was Herbert Spencer. He specifically applied these ideas to the human mind and society. He suggested that the nervous system began as something simple and homogeneous but transformed into something differentiated and very complex (Hergenhahn 2009). He was credited with the introduction of the term intelligence into psychology (Guilford 1967). These theories made it into scientific literature only after naturalists and evolutionists, such as Jean Lamarck and Herbert Spencer, began discussing the altering of species of animals, plants, and humans. With their assistance, Darwin was later able to produce theories on the transformation of animals and humans that would forever change science and influence psychology's views of intelligence.

Darwin was the fifth of six children born to Robert and Susan Darwin. He was described as a "dreamy, grey-eyed, thickset child, intent on his own thoughts behind a shock of brown hair, but warm-hearted and loving for all" (Browne 1995, p. 10). Even from an early age, he was an avid collector of antiques and natural objects. The plan for Darwin was to attend Shrewsbury School, located at the center of Shrewsbury, Shropshire, England, where he grew up. Darwin did not receive this schooling well, stating, "Nothing could have been worse from the development of my mind" (Darwin 1892/2000, p. 9). Despite being busy with subjects such as Greek and Roman history and literature, which were then considered subjects essential to a gentlemanly education, Darwin still chose to use his free time collecting specimens (Stefoff 1996). He was a mediocre student (Berra 2009).

Because Darwin was drawn to natural history and science from a very early age, he and his brother, Erasmus, set up a lab to carry out experiments (Browne 1995). Darwin continued these experiments even after his brother left for Cambridge University (Browne 1995). After leaving Shrewsbury at the age of 16 (Browne 1995), Darwin met up with Erasmus in Edinburgh, Scotland, fully intending on becoming a doctor

J. Rigby, Psy.D. (✉)
Neurology, Learning and Behavior Center, 230 South 500 East, Suite 100, Salt Lake City, UT 84102, USA
e-mail: Jordan@samgoldstein.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_7,

like his father (Browne 1995). He was the third generation of Darwin to study medicine. It was not long before medicine began to look less attractive. When Darwin returned to Scotland to begin his second year of medical school, he found the lectures boring and the horror of surgery without anesthesia too much for him to manage (Berra 2009). He once again turned to natural history and all but ignored his medical lectures to attend lectures on taxidermy and presentations at scientific societies (Berra 2009). Darwin then transferred to Christ's College at Cambridge to study for the church. He was trained to become an Anglican clergyman and graduated 3 years after beginning his studies (Hergenhahn 2009).

Darwin maintained interests in subjects other than the clergy while at Cambridge, such as entomology. This opened the door for networking with professionals in other fields. One such contact would send Darwin down a path that led to the foundation of his theories and influenced his career more than any other (Berra 2009). John Henslow was a botanist who was impressed by Darwin. Henslow was everything Darwin hoped for in a scientific mentor, even to the point of teaching Darwin to use technical instruments (Browne 1995). This mentor-mentee relationship eventually led to a recommendation by Henslow for Darwin to be invited to participate as an unpaid naturalist aboard the ship *Beagle* (Browne 1995). Against his father's objections and an actual list of reasons stating why it was not a good idea, he could not refuse (Berra 2009).

Despite becoming very ill upon setting sail, Darwin considered his voyage on the *Beagle* as "my second life." The setting off of the *Beagle* was "a birthday for the rest of my life" (Darwin 1892/2000, p. 120). This trip was important because it gave him a chance to sharpen his skills of observation and description of people and places. Darwin was under the impression that he would be gone for no more than 3 years depending on the weather (Browne 1995). The mission was to survey South America, take longitude readings, and make geophysical measurements (Berra 2009).

Based on detailed logbooks and journals Darwin kept throughout the entire journey, it is evident that it was his nature to think big and differently than those who first instructed him (Browne 1995). Initially, he thought he might write a book on the geology of the places he was visiting (Berra 2009). He spent weeks at South American destinations such as Brazil, Argentina, and Chile. Much of his work was self-imposed and difficult, but he wished for nothing more than the opportunity to complete it. He enjoyed this work and it was a source of great pleasure and pride to be involved in it (Browne 1995). A key to the excellence of this research was his meticulous approach. He acquired a huge collection of specimens. He was assigned an assistant, Syms Covington, to help skin and clean birds and mammals. They would also sort and pack shells, plants, bones, rocks, and fossils for shipment back to Henslow (Berra 2009).

Darwin became interested in the application of transformation to humans and the perceived ability of human beings to change. This was fundamental to his later conversion to evolutionary theory. He was intent on identifying human beings' place in the natural world (Browne 1995).

After departing from South America, the *Beagle* headed west toward the Galápagos Islands. Darwin visited four of the 16 major islands (Berra 2009). He studied huge tortoises, lizards, sea lions, and 13 species of finch (Hergenhahn 2009). He was especially fascinated with identifying subgroups between specimens from island to island. It was not until the *Beagle* sailed to the Galápagos Islands that he was able to confirm his belief that most animals and plants spread only as far as geographical barriers allowed. With geographical barriers in place, Darwin was able to separate subfamilies of animals based on physical features (Browne 1995).

It was during this voyage that Darwin's idea of becoming a clergyman disappeared. He did not reject God, at least at first, by still giving a spiritual meaning to the natural wonders he saw on the *Beagle* (Browne 1995). The *Beagle* continued on to Tahiti, New Zealand, and Australia (Hergenhahn 2009). The trip ended up being 5 years long. Darwin was eager to return home. When the *Beagle* docked in Falmouth, England, Darwin was one of the first to get off (Browne 1995).

Due to interest among London's naturalists, Darwin believed that his primary scientific focus, following his time on the *Beagle,* should concern the fossils he located. It was not until later that he focused on his animal specimens (Browne 1995). Perhaps a turning point in his research was when he began asking questions regarding the distribution of species among the Galápagos. For Darwin, the possibility of a common ancestry based on specimen's physical resemblances drew his attention from the fossils. He eventually concluded that everything, including mankind, was linked to one ancestral chain through transmutation (Browne 1995).

Darwin's first attempts at evolutionally theory needed refining. His earliest evolutionary ideas were disorderly, scattered, and seemed to ramble. They primarily circled around the question of how transmutation might actually work. Despite being more of a question than a theory at this point, transmutation engulfed Darwin's time and efforts (Browne 1995). At this time, his ideas were different than would be put into print years later. He possessed the pivotal idea of change in living beings and of real links to the ancestors of animals and mankind. However, he had yet to include one piece that could tie it all together, natural selection.

This needed principle to tie everything together came from *An Essay on the Principle of Population* (1798/1914) by Thomas Malthus. An economist, Malthus observed that food supply increased arithmetically, while the human population tended to increase geometrically (Hergenhahn 2009). Furthermore, from Malthus's essay, Darwin recognized that differential death rates gave him a mechanism for explaining the preservation of adaptations, which until then was a mysterious quantity (Browne 1995).

One thing was clear, the Darwin who set sail on the *Beagle* years earlier would likely have had a difficult time formulating these ideas. The collection of his personal ambitions, social background, behavior, desires, and shortcomings contributed to the development of his ideas as much as solid scientific reasoning and research. The range of his intellectual experience was most impressive. His notebooks demonstrated a breadth of ideas and information (Browne 1995). Darwin delayed publication of his theory of evolution for more than 20 years. There is some evidence to suggest that Darwin even intended on having his work published after his death. He wrote his wife Emma a letter that discussed a preference to be dead rather than suffer the controversy of his writings, which he thought would lead him and Emma to social exclusion (Browne 1995). An external motivator could have been responsible for him publishing his theory (Hergenhahn 2009).

Darwin received a letter from Alfred Wallace, who was also influenced by Malthus's essay. In the letter, he described a theory of evolution almost identical to Darwin's. Yet, because of the abundance and meticulousness of data collected by Darwin, the theory of evolution is attributed to him, not Wallace (Browne 1995).

Prior to receiving the letter from Wallace, Darwin had already written a quarter of a million words in his "big book," which he expected would fill three volumes. After the letter, Darwin ceased work on the "big book" and focused on what he called an "abstract" of his species theory (Berra 2009). This would eventually become *On the Origin of Species* (1859). Darwin proposed that natural selection operated on a living being as if it were a statistical necessity, a law of nature stripped of divine influences, relentlessly honing animals, plants, and humans in the struggle for existence. Organisms either adapted or died (Browne 2002). Darwin was not able to use the term "evolution," at that time. The term was mostly used to describe the embryological process of a gradual unfolding of hidden structures. It was the ensuing debate around his published work that changed the word to its modern meaning (Browne 2002). In *On the Origin of Species*, Darwin referred to "descent with modification" (1859, p. 236). He suggested that a struggle for survival of all living organisms is related to the reproductive capacity being greater than the environment can sustain. Among offspring there are infinite individual variations, some of which are more conducive to survival than others. This results in the survival of the fittest; survival results in a natural selection among the offspring of a species. The individual characteristics that are naturally selected are theorized to be more adaptive and have more of a survival value than those that are not passed down through offspring. Fitness is measured by a species' ability to survive and reproduce. Those organisms possessing adaptive features are fit, while those that do not are unfit. Darwin believed that this process just happens without direction or purpose involved. This process is dependent on the environment in which the species exists (Browne 2002).

On the Origin of Species was met with both criticism and praise. Darwin's theory was attributed to the creation of a new science and philosophy. Not until this time had a new branch of human knowledge been due to the fruits of one researcher (Berra 2009). However, Darwin's theory in *On the Origin of Species* was criticized, even to the point that its publication ended some of Darwin's friendships (Berra 2009). It has been suggested by some that Darwin' s theories contributed to Captain Fitzroy, captain of the *Beagle*, taking his own life 6 years after the publication of Darwin's theory because he felt at least partially responsible for its development (Hergenhahn 2009). Despite a mixed welcome, Darwin's publisher, John Murray, received 1,500 orders when it was released (Browne 2002). *On the Origin of Species* went through six editions and has been translated into 29 languages (Berra 2009).

After the publication of *On the Origin of Species,* Darwin continued to publish in spite of a chronic physical ailment. He turned his attention to botany for a time. It was not until Darwin's book *The Descent of Man, and Selection in Relation to Sex* (1871/1896) that Darwin began to build a foundation that humans are also the product of evolution. He suggested that both humans and apes descended from a common ancestor. This was deliberately omitted from *On the Origin of Species*. Also, in *The Descent*, Darwin used the term "evolution" for the first time in the modern sense (Berra 2009). He was interested in "human variation, geographic diversity, facial expressions, moral sensibilities, inheritance, reproductive behavior, and sexual selection" or the entire natural history of mankind (Browne 2002, p. 307).

It was in *The Descent* that Darwin laid out how evolutionary theory explains the intelligence of mankind. He suggested that behaviors commonly associated with modern intelligence developed from the primal instincts of nonhuman ancestors. This led to human intelligence and animal intelligence being differentiated by degree, not by type. Darwin believed that the intellectual abilities of man, and their variations, were inherited. It was to this intellectual ability that mankind owed for its key position in the world (Darwin 1871/1896). Darwin offered examples to support the theory that humans and animals share many cognitive attributes like wonder, curiosity, long-term memory, the ability to pay attention, to imitate the behavior of others, and to reason (Darwin 1871/1896). He suggested that civilized nations would have a tendency to increase in the number and in the standard of the intellectually able. There is an innate tendency toward continued development in mind and body. This leads to the more intelligent genes within the same community being able to succeed in the long run in leaving more offspring, resulting in the inferior genes being counted out of the next generation. This is a form of natural selection (Darwin 1871/1896).

Not all of Darwin's contributions to intelligence theory were positive. He entertained a number of mistaken beliefs by today's standards, one of which involved the intellectual ability of men compared to women. He believed that women were intellectually subordinate to men (Alland 1985). He also suggested that contemporary or primitive people were the link between primates and modern humans and were, therefore, inferior (Hergenhahn 2009).

Late in the year of 1881, Darwin began to suffer from chest pains which progressively got worse. Despite the best efforts of various doctors, he died on April 19, 1882, following a heart attack the night before where he briefly lost consciousness. He was 73 years old. Primarily he was to be buried in a churchyard next to family members near Downe, the small village just outside London where he and Emma had been living for the past 40 years; however, at the request of the Parliament, a funeral was held at Westminster

Abbey. Notable men from many scientific, naturalist, government, and philosophic circles were in attendance (Browne 2002). All of his surviving children attended, as well, but not Emma. His grave is a few feet from the resting place of Isaac Newton and Charles Lyell. Being buried next to others who had such an impact on science is appropriate since Darwin's theory was just as impactful. His ideas are attributed to changing the history of philosophy and psychology by influencing traditional views of human nature. He inspired others to not just look at the mind but behavior for individual differences. This had a direct influence on the introduction of new schools of thought such as functionalism and behaviorism, where individual differences and measurements were emphasized (Hergenhahn 2009).

Francis Galton

Francis Galton was born February 16, 1822, to a wealthy and distinguished English family. He was the youngest of seven children. It was not long before Galton began to demonstrate an intellectual brilliance. By 18 months, he was familiar with the alphabet. He began reading books by 2 years old. At age 5, he was able to recite a Sir Walter Scott poem (Gillham 2001). Galton would impress visitors by reading Shakespeare as early as 6 years old (Fancher 1985). He was instructed by his older sister, 12 years his senior, in French, Greek, and Latin. Galton greatly enjoyed school. Before long he was doing so well that he made head boy, despite being younger than most of the other children. One teacher reported, "the young gentleman is always found studying the abstruse sciences" (Pearson 2011, p. 67). Galton developed a sense of brightness while young. He was motivated to excel in academic competition.

The subject of science, among other things, was important to the family. Galton's mother was the daughter of Erasmus Darwin, who made his name as a poet, physician, and early evolutionary theorist. Charles Darwin was the older half cousin to Galton. Galton's father descended from the founders of the Quaker religion. The family primarily made its money manufacturing guns, but later through investment banking. His father also joined the Church of England (Fancher 1985).

Despite this lineage and a flare for academics, Galton was not really prepared with professional or substantial education, despite his sister's best efforts. His natural inclinations for curiosity and intellectual restlessness were considered distracting and negative qualities when he entered formal British education. For much of his academic career, he was considered mediocre. This did not stop him from viewing his education as a competition among classmates. After finishing his second year of medical training at 18 years old, he wholeheartedly believed that he was going to finish first on the final medical examination. He only won a Certificate of Honor instead of the top prize (Fancher 1985). Perhaps too painful to admit the defeat, Galton incorrectly remembered finishing first on this examination in his autobiography written 70 years later (Galton 1908). Throughout his career at Cambridge, he did just well enough to keep his hopes alive to win honors in mathematics. It was during this time he became interested in the benefits of examinations themselves. The idea of an instrument that was able to distinguish students at the top from the rest of the group was an untapped resource as far as Galton was concerned (Fancher 1985). A measurement that would identify students who stood above others academically would be useful. His desire that exams would place him in a superior category in the area of mathematics was never realized. The collection of his work earned him the equivalent of a B + average at one of the most competitive universities in the world (Fancher 1985). This was not the outcome that Galton desired or anticipated. This crushed his dream to be a top *Wrangler*, a title reserved for students at Cambridge who earned the top scores on each year's mathematics honors examination. Shortly following this realization, he withdrew from future exams and studying mathematics altogether. This was followed by what some consider an emotional breakdown (Fancher 1985).

The last year of Galton's time at Cambridge was filled with dinner parties, balls, and wine parties. By this time he was an attractive young man with blond hair, high forehead, pale blue eyes, and v-shaped mouth (Gillham 2001). The end of Galton's academic career was not what he nor his disappointed family had envisioned. He was only able to earn a "poll" or ordinary degree, much like his cousin Darwin. With his formal academic career over, he failed to prove himself the genius both he and his family envisioned.

Galton earned a substantial inheritance when his father passed away in 1844. After several years drifting and living on his inheritance, he consulted a professional phrenologist for a personality profile that would steer him in the right direction (Gillham 2001). The professional assessed that Galton's skull, with other information, suggested men with his type of intellect do not distinguish themselves in universities. Galton turned his attention to Africa. He organized a major expedition through South-West Africa that took him away for a couple of years. This would lead to the creation of one of the first maps of the area. Similar to Darwin's voyage on the *Beagle*, Galton's travels greatly contributed to the formation of his theories (Fancher 1985).

Upon returning home to England from Africa, Galton was different. He was matured and experienced by adversity (Gillham 2001). This change motivated him to work on what he referred to as the "human side of geography" while also working on informative books for travelers (Pearson 2011, p. 129). He was impressed by the diversity between different African groups. This awareness of ethnic diversity produced several major ideas which today are at the heart of the nature-nurture and IQ controversies (Fancher 1985).

Despite his successes, a question continued to plague Galton. Coming from a family of privilege and having so many academic opportunities, he had trouble explaining how his environment could not produce the genius most thought he would become. He was told by the phrenologist that despite these advantages, it was insufficient innate ability that he lacked. He would later agree with this theory in his book *Hereditary Genius* (1869), when he drew a resemblance between intellectual capacity and physical capacity. He argued that with physical training, physical improvement could be expected. However, there is a genetic limit or ceiling when improvement stops, no matter the time or effort devoted to training. Galton said the same about intellectual ability:

> This is precisely analogous to the experience that every student has had of the working of his mental power. The eager boy, when he first goes to school and confronts intellectual difficulties, is astonished at his progress. He glories in his newly developed mental grip and growing capacity for application, and, it may be, fondly believes it to be within his reach to become one of the heroes who have left their mark upon the history of the world. The years go by: he competes in the examinations of school and college, over and over again with his fellows, and soon finds his place among them. He knows that he can beat such and such of his competitors; that there are some with whom he runs on equal terms, and others whose intellectual feats he cannot even approach. (p. 13)

In Galton's history and experience, he observed the differences between those who had similar environments. This appeared to contribute to his not completely eliminating the influence of environment on intellectual development, but heavily favoring the role of innate abilities (Fancher 1985).

Galton was also influenced, like much of the scientific world, by Darwin's *On the Origin of Species.* Despite having no real interest in biology, at least not to the degree of his cousin, he was very interested in evolutionary theory as it applied to the diversity of humans. This was the root of his own ideas that took on great significance and occupied the majority of his life's work (Fancher 1985). Darwin's writing "made a marked epoch" in how Galton understood the mind (Galton 1908, p. 287).

Galton was excited by Darwin's idea that while unfavorable variations in breeding populations were not passed on, favorable variations would increase in frequency over generations. Darwin applied this theory to physical characteristics. Using this same theory, Galton thought that perhaps psychological differences were inheritable and based on small variations in the brain and nervous system. Adding to his already established belief that ability was innate, the idea of inheritability took on a whole new meaning with practical applications (Fancher 1985). At the time, readers were just beginning to consider the unconventional thought of physical and some mental traits being inherited from generation to generation in animals. Many were unprepared to consider that theory being applied to humans (Gillham 2001). What Galton was proposing was that human talent and character changed similar to the domestic animals and cultivated plants discussed by Darwin (Gillham 2001). Humans were subject to natural selection. Galton began to see the advantages in this process by enhancing qualities through selective breeding. Galton first made these thoughts public in a two-part article published in *Macmillan's Magazine* in 1865. These series of papers were the precursors to Galton's 1869 book *Hereditary Genius.* Darwin tended to agree with Galton stating, "I am inclined to agree with Francis Galton in believing that education and environment produce only a small effect on the mind of anyone, and that most of our equalities are innate" (Gillham 2001, p. 155).

In *Hereditary Genius* Galton compiled a sample of those whose reputation suggested they had shown unusual talent in their lives. He did not include those who were considered brilliant due to ancestry, like hereditary aristocracy (Fancher 1985). He used two systems to classify the men's ability. First he classified men by reputation, then by their performance on examinations. However, he was unable to apply the examination scores to a normal distribution. So this was abandoned, and the first classification method was the only inclusion criteria used (Gillham 2001). He would use men's reputations to measure heritable ability. By utilizing statistical means to analyze the data, a somewhat novel approach for the subject at the time, he determined that his sample represented a very small percentage of the normal population, about one person in 4,000 (Galton 1869). After establishing the infrequency associated with his sample in the general population, he cross-checked their lineage and found 10 % of them had at least one close relative in the sample. Immediate family was much more common than distant relatives. The sample was full of fathers, sons, and brothers. Talents were also not evenly distributed among the sample. There was a propensity for an unusual talent to run in a family (Galton 1869). Again, expanding on Darwin, Galton determined that these heredity similarities could be put into the same selection process as physical similarities. His reported results, although very general and not conclusive, were consistent with his hypothesis that differences in ability are inherited. Galton understood the limitations to his research alluding to there being no advantage as beneficial as a good education (Fancher 1985). He suggested behind heredity, families shared environmental circumstances. Given his own academic experience, it is understandable how he would minimize the importance of education.

In an effort to establish empirical support for his ideas, Galton presented two research techniques still used today. First, he utilized the role of adoptive families who have shared environments but not genetics. Critics at the time had issues related to his sample size being truly representative. Despite these limitations, adoptive families were useful in identifying hereditary similarities (Fancher 1985). The other technique employed involved twins. He understood the potential both dizygotic twins and monozygotic twins have in understanding the environment's versus nature's influence on development. With his data in hand, he applied the bell curve as an evaluative technique.

With the ground work laid by *Hereditary Genius,* Galton could move forward with his eugenics program. He was passionate about this. It became a matter of moral and civic obligation. He made efforts to improve the intellectual standard of humans, not through educational or environmental reform but genetics (Fancher 1985). His mission to accomplish this was twofold: first the spread of intellectually and psychologically superior humans through reproduction of offspring; second the establishment of laws and cultural customs to encourage superior humans to flourish at a faster rate than those deemed inferior. In doing so, these favorable genes would be more numerous in society and would increase in quality through natural selection. Just before his death he finished a novel, *Kantsaywhere,* about a eugenic utopia. Although in name only, Galton's legacy of eugenics will be forever tied to such programs as the Nazis encouraging racially appropriate women to have larger families to increase the number of desirable individuals. This cruel practice of Nazi policies, ultimately resulting in the "final solution" and the holocaust, led to a total revolt counter to eugenics after World War II (Gillham 2001). Although Galton created the term "eugenics" and campaigned for it, he was simply generalizing Darwin's theory of evolution through natural selection to mankind. He would have likely been shocked to hear that only 20 years after his death, the name of eugenics was being used for forcible sterilization and murder (Gillham 2001).

It was Galton's intention to identify prominent or eminent individuals so that they could pass on their superior genes to offspring. These individuals were prominent due to their numerous positive contributions to society. However, the timing of this societal status was customarily specified after prime childbearing age. This was a problem for Galton. How was he going to identify those prominent individuals before this prime age passed? A measure was needed to recognize a young adult's ability. After being measured, this natural ability could be used as a predictor of who would make eminent contributions later in life. Those deemed favorable to make these types of contributions would then be encouraged to intermarry and produce offspring (Fancher 1985). Galton was holding on to his basic belief that there is a measurable difference between individual's brains and nervous systems that can be differentiated. He went on to design a series of measures for sensory perception, reaction time, and physical energy (Gillham 2001). The goal was to identify differences in neurological efficiency or what he would refer to as natural intelligence.

In order to gather data, he set up, equipped, and maintained the Anthropometric Laboratory at his own expense. It was Galton's desire "to compile a list of instruments suitable for the outfit of an Anthropometric Laboratory especially those for testing and measuring the efficiency of the various mental and bodily powers" (Gillham 2001, p. 211). Galton even wrote to the leading psychologists in England to allow him to display their names to what he had in mind. However, much of the experimental psychology and sensory perception work was being done in Germany by men like Herman Helmholtz and Wilhelm Wundt (Gillham 2001). Demographics were gathered, and participants in the lab took various tests in sequential fashion over the period of about a half hour. They were measured with varying devices assessing "Keenness of Sight and of Hearing; Colour Sense, Judgment of Eye; Breathing Power, Reaction Time; Strength of Pull and Squeeze; Force of Blow; Span of Arms; Height, both standing and sitting; and Weight" (Gillham 2001, p. 213). To measure the strength of a blow, for example, Galton used a rod made of fir protruding out of a tub so that it could move freely. A buffer to prevent the participants from hurting his or her hand after smashing their fist down was attached. Following an impact, a pointer attached to the rod measured the distance the rod traveled in the tube (Gillham 2001). Primarily, Galton avoided measuring head size. He felt that women's heads were too difficult to measure "on account of their bonnets, and the bulk of their hair" (Gillham 2001, p. 213). This practice was later included in the battery of measurements.

The Anthropometric Laboratory was a big success despite many of the tests never really working as intended. Many failed to correlate with any independent signs of accomplishment or intelligence (Fancher 1985). But, by its end in 1885, statistics had been collected on 9,337 people, each measured in 17 different ways (Gillham 2001). Each person was paid three pence to participate in the experience (Fancher 1985). Participants even returned for repeated measurements. Before its closing Galton began collecting fingerprints. This was a newly acquired interest that turned into a major branch of his work. He also studied mental imagery as a quantifiable measure of the human mind. Based on one's "imagery ability," they could be compared to others. From this came the first word association tests. It is not known whether or not this influenced Sigmund Freud, but Galton was utilizing concepts of psychoanalysis (Hergenhahn 2009). Galton collected an enormous amount of data that required analyzing. In doing so, he developed his concept of regressing, invented the correlation coefficient, and proposed one of his most important and influential books, *Natural Inheritance* (Galton 1894). The Anthropometric Laboratory had also stirred up an interest in testing ability as a new area of research. As a result, mental heredity began to become commonly known and accepted (Fancher 1985).

Sir Francis Galton died on January 17, 1911, after approximately 5 years of failing physically but remaining mentally whole. Throughout his life, he was involved in exploration and travel writing. He was largely responsible for the development of fingerprinting as a forensic method. Galton's contributions to mental imagery influenced psychology, along with methods in pedigree analysis and twin studies. He introduced the statistical method of correlation and regression, resulting in the use of quantitative methods into social sciences (Gillham 2001).

Jean-Martin Charcot

Born in Paris in 1825, Jean-Martin Charcot was raised in a lively section of town that is now considered the Right Bank, north of the Le Seine River. In Charcot's time this was between the central city and the countryside. He was the eldest of four sons. He did not go far to attend grammar school at the Pension Sabatier. Among a very standardized curriculum for all schools that included a national examination at the end of each school year, he studied Greek and Latin in preparation to transfer to the Collège Royal de Bourbon after graduation. Although he took the prerequisites to enter medical school, such as Latin, he was not sure medicine was his calling. He was torn between medicine and art as a career path. He eventually chose medicine and began his training at the Paris Medical School in 1843 (Goetz et al. 1995).

So began Charcot's career in medicine. He was described at that time as a thin, pale man with long black hair tossed back and a small black mustache who did not get out to socialize much. His program was filled with students from across the globe: British, American, German, and Austrian. This program involved a fixed yearly curriculum, forfeiting traditional grades for a series of year-end tests to graduate. Clinical competency or practical skills in medicine were not weighed heavily on these final examinations. Theoretically, it was possible to pass all medical examinations without actually interacting with a patient. Hospitals were used for teaching and part of the curriculum, but much of what was offered was only taken advantage of by the more assertive students. One year of clinical rotations was required the year Charcot began his training (Goetz et al. 1995).

This was not sufficient for Charcot. He was interested in pursuing an internship. At the time, this position was optional and highly competitive. Only the elite students were selected for this additional clinical experience. It took Charcot 2 years before he was accepted into this special group. After his first application, he was given the status of a temporary intern before becoming one of six interns. Internship consisted of four yearlong rotations at different hospitals under different chiefs (Goetz et al. 1995).

Two of these supervisors had a particularly great influence on Charcot, P. A. Piorry and Pierre Rayer. Piorry did not study the nervous system, but his emphasis on anatomical localization and nosology influenced Charcot. Piorry also taught Charcot the practical use of microscopy in medicine. The most influential supervisor in Charcot's early career was Pierre Rayer. He was a mentor to Charcot. As a result of Charcot not having any familial connections into the closed scientific and social circles of medicine, Rayer befriended him and aided in networking him into those closed groups. This was a practice not lost on Charcot, as later in his career he protected and nurtured the careers of his protégés one of which was Sigmund Freud. Freud was interested in studying children's brains and was welcomed by Charcot. Freud offered to translate text into German and was quickly accepted into Charcot's good graces.

Charcot found himself in a very advantageous situation. He was being trained by important and influential physicians in what was the medical capital of the world. He was intelligent and possessed a unique capacity to work. But what was perhaps a catalyst to this "perfect storm" that perpetuated him into international success was a determination to publicize his ideas. He was established in publications very early and maintained his scientific involvement throughout his career.

Charcot took 9 years to graduate. This was the typical length taken before completion of medical school (Goetz et al. 1995). He was awarded his diploma in 1853. He received an exceptional grade on his final examination, which was a thesis on rheumatoid arthritis. He collected his data where he completed his internship, at the Salpêtrière. Overall, his grades ranged from "rather good" to "very good." There is no evidence to suggest that Charcot had to retake or failed any courses, a common occurrence for many average students who did not pursue internships.

Following graduation Charcot accepted a position as a head of a clinic. There he worked for two years under his old supervisor, Piorry. This position was not highly sought after or lucrative but was a natural step-up in the academic hierarchy. Following his clinic position he became a hospital doctor for the next 4 years. This may have been disappointing due to few opportunities to study patients. He primarily handled outpatient triage in the central hospital admitting office. Despite his position, Charcot worked toward building a lasting reputation (Goetz et al. 1995).

His ultimate goal was to secure a position at a Paris hospital with the medical faculty. He set out to do this by occupying his free time with building a private practice. A private practice allowed Charcot to develop a positive reputation and helped financially. He also continued to publish and conduct research within medical scientific societies. He managed to make significant contributions involving the subjects of pathology and critical reviews. Even these earliest publications demonstrated his abilities as a capable, wide-ranging, and discerning consumer of the available literature. The last thing he did to move closer to his goal was to teach both outside and within the medical program.

It was Charcot's desire to return to the Salpêtrière. This is where he had begun to study the nervous system and where he would become known all over the world. Most physicians at the time would rotate from hospital to hospital. Charcot broke that pattern and established himself at the Salpêtrière for the majority of his career. Given his work upon starting at the Salpêtrière, he may have been motivated by the opportunity to conduct original research as he made this an emphasis at the very beginning. He focused on pathology, diseases of the nervous system, and chronic rheumatism above the rest (Goetz et al. 1995).

Over the course of many years, Charcot was able to overcome that which held him back from a failed bid for a medical professorship in internal pathology. A vote took place, the results of which were overwhelmingly not in Charcot's favor. He received only a handful of votes for one of two positions. However, the period following his defeat has been described as advantageous due to his efforts for research and teaching at the Salpêtrière. This additional experience led to a convincing victory years later. His research, specifically focusing on pathology of the nervous system, and his style of teaching made him a very strong candidate. He became chair of pathological anatomy at the Paris Medical School in 1873, but still continued his work at the Salpêtrière. Obtaining professorship when he did, not earlier or later, clearly contributed to his career in neurology as he was later appointed the professorship for the study of disease of the nervous system in 1882. This allowed him to focus on neurological issues.

Charcot's interests progressed from general medicine among a geriatric population to chronic illnesses afflicting all ages. At the Salpêtrière Charcot was responsible for groups classified as rheumatological and neurological. In an effort to explain the medical findings in arthritic and neurological populations, Charcot developed an anatomo-clinical method of research. This would become the framework of which all of Charcot's anatomical studies in neurology were based.

This linked clinical signs with anatomical lesions. Charcot utilized longitudinal observation to meticulously document clinical signs. This data was then cross-examined with the brain and spinal cord following an autopsy. He hoped to establish anatomo-clinical correlations by combining clinical and anatomical data. Through sound research he hoped to locate specific lesions which would correspond to specific clinical signs (Goetz et al. 1995). This was the foundation for a new anatomy-based classification of neurological pathology. Perhaps the best-developed example of this method was Charcot's work with motor system degenerative disorders, specifically amyotrophic lateral sclerosis. These studies led to the international description of amyotrophic lateral sclerosis as Charcot's disease, now commonly known as Lou Gehrig's disease. Also, numerous stroke syndromes and the relation of specific signs of multiple sclerosis to specific lesions are the product of the anatomo-clinical method. This discipline pushed neurologists toward the notion of cortical localization theory, suggesting the brain controlled specific motor, sensory, and language functions (Goetz 2010).

Charcot became increasingly interested in studying hysteria and hypnosis. He employed the help of volunteer researchers, one of whom was Alfred Binet. Questions being raised concerning these two conditions were motivation and unconscious psychological effects. He displayed a considerable amount of interest in the psychology side of hysteria, more than most of his colleagues in medicine. Even psychiatrists of the day were committed to somatic and hereditary assumptions about mental diseases and had very little interest in the possibility of a psychological etiology. What sparked Charcot's interest in hysteria was that the symptoms so closely imitated known neurological pathology but were absent from a neurological or organic cause. After an exhaustive neurological exam exposed no pathology, those suffering from hysteria still presented with anesthesias, amnesias, or paralyses. He also noted a traumatic event prior to the onset of symptoms in several of the hysteric patients. Although not physiologically or neurologically damaging, Charcot suggested that these traumatic events sparked ideas that resulted in hysterical symptomology (Hergenhahn 2009). At the time, these symptoms were often disregarded as malingering. The majority of physicians were suspicious. It was Charcot's thought that this directly related to hypnosis given that symptoms of hysteria could be reproduced so readily in a hypnotic state. During a hypnotic trance Charcot could induce hysterical body contractures. This gave evidence to his theory that hypnosis could expose the true mechanisms of hysteria (Fancher 1985). Being components of the same neurological state, the capacity to be hypnotized and the diagnoses of hysteria were linked (Goetz et al. 1995). This was something Charcot advocated, even near the end of this life.

Charcot, with Binet's help, wrote case studies that would illustrate these theories and highlight the complexity of real people. Charcot's theories complicated the established distinction between biological and psychological investigation, allowing for further research focusing on intelligence and its relationship with neurophysiological and genetic factors. His theories did not go unchallenged. Many claimed that any suggestible person could be hypnotized, independent of hysteria. This thinking made hypnosis available as an intervention for diagnosis outside of hysteria. Many clinicians interested in exploring hypnotism, including Freud, tended to agree. As time passed, Charcot's views on hysteria and its relationship to hypnotism became increasingly difficult to defend as new insights on the subject emerged (Goetz et al. 1995).

Another contribution by Charcot was with diseases of the nervous system. Charcot made an argument that as an organism evolved to become more complex, the nervous system took on an ever progressively dominant role over all other functions along with a close connection anatomically (Goetz et al. 1995).

Although Charcot may have held others' health and well-being in the highest regard, he did not do the same for his own. He sustained a high degree of intensity, both personally and professionally, that likely attributed to his death. He also ate heavily and smoked endlessly. He had personal interests, but he seemed to always incorporate some type of medical element. Despite serious cardiac problems, Charcot never cut back his work. He died at age 67 of pulmonary edema early in the morning on August 16, 1893, at an inn near the Morvan area of France. Just before his death he wrote a letter to his wife stating that he had not been feeling well. His death was news that quickly traveled worldwide via medical journals and newspapers. By the time of his death, Charcot was a well-respected contributor to the field of neurology. Although much of his work with hysteria was imperfect and most was replaced by others such as Freud, his observations contributed to the concept of treatment for psychosomatic conditions and the association of psychology into the manifestation of nervous illness (Goetz et al. 1995). His anatomo-clinical model was unique and was being developed as a discipline all over the world. Charcot's unique personality and history directly influenced its development and is still used to distinguish neurology from other medical disciplines. The classification system that Charcot created for defining the brain and nervous system and muscle diseases to this day can be found in modern neurological classifications. For this reason, he will forever be considered the "founder of modern neurology."

References

Alland, A., Jr. (1985). *Human nature: Darwin's view*. New York: Columbia University Press.

Berra, T. M. (2009). *Charles Darwin: The concise story of an extraordinary man*. Baltimore: The Johns Hopkins University Press.

Browne, J. (1995). *Charles Darwin: Voyaging*. Princeton: Princeton University Press.

Browne, J. (2002). *Charles Darwin: The power of place*. Princeton: Princeton University Press.

Darwin, C. (1859). *The origin of species: By means of natural selection or the preservation of favoured races in the struggle for life*. John Murray. London: England, doi:10.1037/13681-000

Darwin, C. (1871/1896). *The decent of man and selection in relation to sex*. New York: D. Appleton and Company.

Darwin, C. (1892/2000). *The autobiography of Charles Darwin*. Amherst: Prometheus Books

Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York: W.W. Norton & Company.

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences*. New York: D. Appleton and Company.

Galton, F. (1894). *Natural inheritance*. New York: Macmillan.

Galton, F. (1908). *Memories of my life*. London: Methuen & Co.

Gillham, N. W. (2001). *A life of Sir Francis Galton: From African exploration to the birth of eugenics*. New York: Oxford University Press.

Goetz, C. (2010). Jean-Martin Charcot and the anatomo-clinical method of neurology (Chapter 15). In *Handbook of clinical neurology* (pp. 95203–95212). doi:10.1016/S0072-9752(08)02115-5.

Goetz, C. G., Bonduelle, M., & Gelfand, T. (1995). *Charcot: Constructing neurology*. New York, NY: Oxford University Press.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Hergenhahn, B. R. (2009). *An introduction to the history of psychology* (6th ed.). Belmont: Wadsworth/ Thomson Learning.

Malthus, T. (1914). *Essay on the principle of population*. New York: Dutton. (original work published 1798).

Pearson, K. (2011). *The life, letters and labours of Francis Galton: Correlation, personal identification and eugenics*. New York: Cambridge University Press.

Stefoff, R. (1996). *Charles Darwin: And the evolution revolution*. New York: Oxford University Press.

8 Social Competition and the Evolution of Fluid Intelligence

David C. Geary

The hallmark of human competence is the ability to modify evolved brain and cognitive systems to create evolutionarily novel abilities that in turn undergird much of culture as we experience it today. The modification of these primary, evolved systems and the cross-generational accumulation of these advances have resulted in writing systems, mathematics, science, a myriad of technological advances, as well as literature, art, and many other features of the modern world. General fluid intelligence is critical to the creation and learning of these non-evolved abilities and represents an evolved system itself (Geary 2005). Proposals regarding how intelligence evolved emerged soon after Darwin's (1859) *On the Origin of Species* (Darwin 1871; Huxley 1863) and in many respects anticipated the central argument of modern-day theorists (Alexander 1989; Ash and Gallup 2007; Dunbar 1998; Flinn et al. 2005; Geary 2005; Kaplan et al. 2000; Mithen 1996). The central theme is that intelligence represents the ability to anticipate and predict variation and novelty that arise within lifetimes (as contrasted with cross-generational change) and to devise strategies to cope with this novelty. Theorists differ largely on whether the key driver of novelty was climatic change, the nuances of hunting other species, or from the dynamics of human cooperation and competition.

It is of course possible that all of these factors contributed, perhaps at different points during hominid evolution. The key question concerns the selection pressure that was central to the evolution of these competencies, following the emergence of the genus *Homo* and especially the emergence of *H. sapiens*. I first provide respective reviews of research on hominid brain evolution and psychometric and cognitive neuroscientific research on general intelligence. These reviews provide the background needed to appreciate the substantive changes that have occurred in the brain and the mind during human evolution and to understand the component competencies that define intelligence. I close with an attempt to integrate these findings with models of climatic, ecological, and social selection pressures and within an integrative model of human brain, cognitive, and psychological evolution (see Geary 2005).

Evolution of the Hominid Brain

Brain Volume and Organization

The mean brain volume estimates for key species of *Homo* and the predecessor genus *Australopithecus* are shown in the top portion of Fig. 8.1 (Asfaw et al. 1999; Falk et al. 2000;

D.C. Geary, Ph.D. (✉)
Department of Psychological Sciences, Interdisciplinary Neuroscience Program, University of Missouri, 210 McAlester Hall, Columbia, MO 65211-2500, USA
e-mail: GearyD@Missouri.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_8,

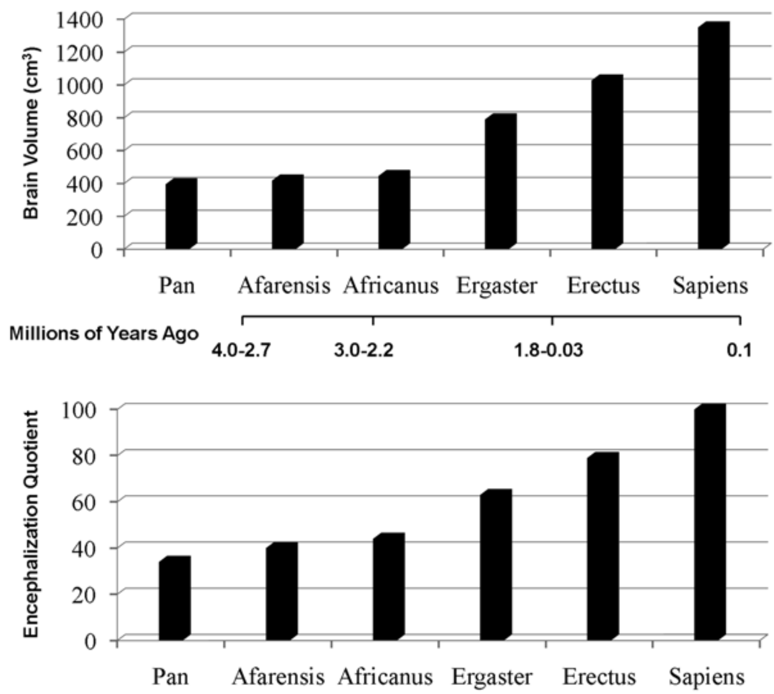


Fig. 8.1 The *top* portion presents the estimated brain volume for chimpanzees (*Pan troglodytes*) and key species of australopithecines (*A. afarensis*, *A. africanus*, *A. garhi*) and Homo (*H. habilis*, *H. ergaster*, *H. erectus*, *H. sapiens*). The *middle* portion presents estimated time of existence of these species; *H. ergaster and H. erectus* are may be earlier and later forms of the same species. The *bottom* portion presents estimated encephalization quotients (EQ) for these species as a percentage of the EQ of modern humans (*H. sapiens*) (Adapted from Geary 2010, p. 136. Copyright 2010 by American Psychological Association)

Holloway 1973; McHenry 1994; Tobias 1987; Wood and Collard 1999). The australopithecines (*A. afarensis*, *A. africanus*, and *A. garhi*) had a modestly larger brain volume than modern-day chimpanzees (*Pan troglodytes*). Further increases in brain volume are evident in *H. habilis* (Tobias 1987) and with the emergence of *H. erectus* to modern humans (McHenry 1994; Wood and Collard 1999). Casts of the fossilized crania provide impressions of the inner surface of the brain and suggest that the australopithecine brain, including the shape of the posterior portions of the left hemisphere and anterior portions of the right hemisphere, differed from that of chimpanzees (Holloway and Kimbel 1986). The visual area of the left hemisphere is smaller than expected based on body and brain size, and the left parietal area (e.g., area 39) and right frontal pole (area 10, Fig. 8.2) area are larger than expected (Holloway and de al Coste-Lareymondie 1982). Expansion of these areas is potentially important, because as they are implicated in studies of fluid intelligence, working memory, and self-awareness (Jung and Haier 2007; Levine 1999; Tulving 2002). The area of the frontal lobe involved in human speech and gesture (i.e., Broca's area) appears to have expanded with the emergence of *H. habilis* and had architecture similar to that of modern humans.

The threefold increase in brain volume from the australopithecines to modern humans is of course substantial but potentially misleading. The absolute size of the brain increases with increases in overall body size and thus confounds cross-species comparisons. An adult male *A. africanus* likely weighed 30 % less than a modern adult human male, suggesting some proportion of the difference in brain volume is due to differences in overall body size (McHenry 1994). The encephalization quotient (EQ) is an often-used

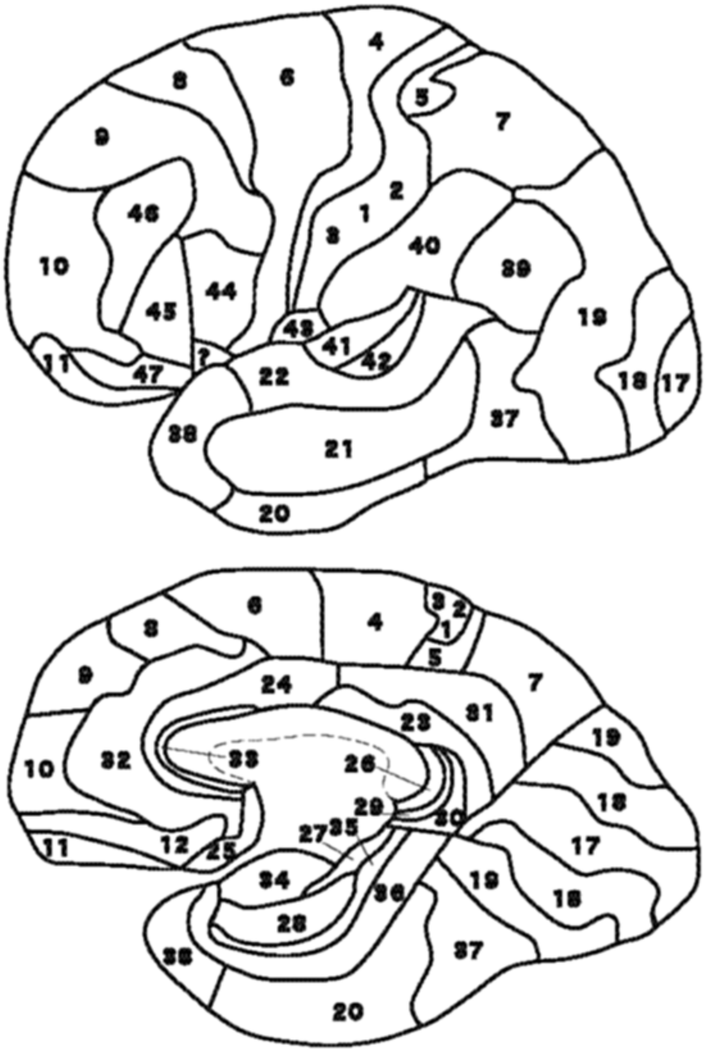


Fig. 8.2 Cartoon maps of Brodmann's (1909) areas of the human neocortex. The *top* section is a lateral (*outer*) view of the cortex, whereas the *bottom* section is a medial (*center*) view. Many of these areas can be subdivided into specialized subregions that may process different forms of information. Generally, areas 1, 2, 3, 5, 31, and 43 are part of the parietal cortex and support a variety of functions including sense of body position, attention, and spatial competencies; areas 17, 18, and 19 are part of the occipital cortex and support simple and complex visual perception; areas 22, 41, 42, and subregions of areas 40 and 38 are part of the temporal cortex and support simple and complex auditory and speech perception; areas 20, 21, 26–28, 34–37, and 52 are also part of the temporal lobe but support a variety of complex visual competencies; areas 4, 6, and 8 are involved in complex motor movements and are part of the frontal cortex; Area 44 and subregions of area 45 are involved in speech generation and are part of the frontal cortex; areas 9, 10, 11, 25, 46, 47, and subregions of 45 are part of the prefrontal cortex and support behavioral control, executive functions, and many complex social competencies; areas 23, 24, 30, (parts of 31), 32, and 33 are part of the cingulate and support attentional and emotional functions

measure to control for the relation between brain and body size; EQ indexes brain size relative to that of a mammal of the same body weight (Jerison 1973), with an EQ of 1.0 indicating an average brain size.

The EQ of chimpanzees is about 2.0 and about 6.0 for modern humans (McHenry 1994). The bottom portion of Fig. 8.1 shows EQ estimates that are expressed in terms of a percentage of that of modern humans, and thus although the EQ of chimpanzees is twice that of the typical mammal, it is estimated to be only 34 % of that of modern humans. The EQ of australopithecines is greater than that of chimpanzees but less than 50 % that of modern humans, with large increases in EQ evident with the emergence of *H. ergaster* and *H. erectus*. Following the emergence of *H. ergaster*, there was a period of little change in EQ for about 1.2 million years ago (Ma), followed by a modest 12 % increase from 500,000 to 400,000 years ago, and then very rapid increases until about 20,000 years ago, and followed by a 3 % to 4 % decline (Holloway 1996).

Climatic, Ecological, and Social Selection Pressures

Climatic and geological changes are common and can result in variation in temperature, rainfall, and other ecological conditions that in turn can influence the mix of vegetation, woodland, and other factors that can have dramatic evolutionary effects on species. Vrba (1995) argued that significant glaciation between 2.8 and 2.5 million years ago resulted in decreased temperature and rainfall in Africa and substantial evolutionary changes in many species, including hominids. Evolutionary responses to decreasing temperature often include increases in physical size to enable retention of body heat and a prolonged period of physical development to enable growth of a larger body. Vrba proposed that these physical adaptations resulted in an accompanying change in brain volume and EQ: "The conclusion is inescapable that hominine encephalization in the latest Pliocene started a new trend, of higher evolutionary rates than before" (Vrba 1995, p. 406).

The implication is that the increase in EQ with the emergence of *Homo* was not the direct result of ecological or social pressures but rather an incidental effect of physical adaptations to climate change. Evidence for Vrba's (1995) hypothesis is mixed. The emergence of various species of australopithecine 3.0 to 2.0 MYA may have been driven, at least in part, by climatic and ecological changes in Africa during this epoch, but does not appear to be the primary evolutionary pressure since that time (Elton et al. 2001; Falk et al. 2000). Other theories have focused on the benefits to being able to predict seasonal changes and prepare for climatic variation (Potts 1998). Some studies do find a correlation between estimates of variability in global temperature and changes in brain volume and EQ from *H. habilis* to early modern humans (Ash and Gallup 2007). Other studies, however, find a smaller or no relation once other factors are controlled (Bailey and Geary 2009; Pearce et al. 2013).

Ecological models of hominid brain evolution have focused on improvements in the ability to extract biological resources from the ecology and through this reduce premature mortality. The basic idea is supported by comparative studies of nonhuman species: Species with complex foraging or predatory demands have a larger brain volume and higher EQ values than related species with less complex foraging or predatory demands (Barton 1996). The basic argument is that hominids evolved into super predators (Wrangham et al. 1999). Humans in traditional societies are indeed highly efficient at extracting life-supporting resources from the ecology through hunting and foraging (Kaplan et al. 2000). If this ability to extract and process (e.g., through cooking) biological resources was the driving force in hominid brain evolution, then there should be evidence of social, behavioral, and other adaptations that allowed evolving hominids to forage and hunt in increasingly sophisticated ways.

There is some evidence that *A. garhi* constructed stone tools, including tools used to cut and process meat (Semaw et al. 2003). Moreover, features of tooth morphology and behavioral adaptation suggest that australopithecines were able to eat a wider range of foods than their predecessors and thus were able to exploit a wider range of ecologies. Wrangham et al. (1999) argued that *H. erectus* used fire for cooking, which enables the use of an even wider range of plant and animal species as foods. Foley and others have demonstrated that tools related to hunting and foraging (e.g., digging sticks) have become increasingly sophisticated since the appearance of *A. afarensis* (Foley 1987; Foley and Lahr 1997). *H. habilis* appears to have used simple stone tools, with increases in the complexity of these tools and their wide geographic distribution coinciding with the emergence and migration patterns of *H. erectus*. These advances continued with the appearance of early modern humans, with the most complex stone tools found in archeological sites dating less than 50,000 years ago (Foley and Lahr 1997). The pattern of technological advances in tools and their utility in hunting and foraging is consistent with ecological models of human brain evolution (e.g., Kaplan et al. 2000).

Patterns of human migration and subsequent mass extinctions of other species are consistent with the evolution of a super predator. Wallace noted as much in 1911 (p. 264) – "the rapidity of the extinction of so many large Mammalia is actually due to man's agency." Decades later, Martin (1973) presented systematic evidence that indeed mass extinctions of many species of large prey were evident in Africa about 50,000 years ago, and later mass extinctions occurred in Australia, Asia, America, and New Zealand after the migration of humans into these regions (e.g., Alroy 2001), although some of these extinctions appear to have been due to climatic or other ecological changes. In any case, declines in many populations of large fish and mammals have been directly linked to human hunting and fishing (e.g., Hsieh et al. 2006). Consistent with these patterns is evidence that the evolutionary increase in brain volume was associated with a corresponding decrease in the mass of the metabolically expensive gastrointestinal tract (Aiello and Wheeler 1995). Reduction in the size of the gastrointestinal tract requires change from a low-quality (e.g., plants) to a high-quality (e.g., fruits, meat) diet, as appears to have been the case during hominid evolution. It is not that reduction in the gastrointestinal tract directly caused the evolutionary expansion of brain volume and EQ but rather released a significant metabolic constraint, creating opportunity for evolutionary change.

The gradual improvement of hunting and foraging abilities, especially after the emergence of *Homo*, resulted in ecological dominance and set the stage for a within-species evolutionary arms race (Alexander 1989; Flinn et al. 2005). Once ecological dominance was achieved, an evolutionary Rubicon was crossed:

> the ecological dominance of evolving humans diminished the effects of 'extrinsic' forces of natural selection such that within-species competition became the principle 'hostile force of nature' guiding the long-term evolution of behavioral capacities, traits, and tendencies (Alexander 1989, p. 458).

Ecological dominance manifests as the ability to efficiently extract biological resources from the ecology, as described above, and the ability to manipulate the ecology in ways – building of shelters, use of fire, clothing, and so forth – that reduce mortality risks and supports subsequent population expansions (Hill et al. 2001; Kaplan et al. 2000). The lower mortality that would follow from ecological dominance would have necessarily resulted in population increases, which would result in pressures for migration to new regions, as happened with *H. erectus*. Population increases also carry the risk of expansion beyond the carrying capacity of the ecology, which results in fewer available resources per capita and an ensuing increase in the competition for these diminishing resources. The pattern was of course described by Malthus (1798) and confirmed in subsequent studies of population dynamics in developing Europe and developing nations today (e.g., Hed 1987; United Nations 1985). The end result is often a population crash that disproportionately affects individuals who are economically poor and of lower social status. The inverse relation between social status, resource control, and mortality risk creates a never-ending cycle whereby Darwin's and Wallace's (1858, p. 54) conceptualization of natural selection as a "struggle for existence" becomes in addition a *struggle with other human beings for control* of the resources that support life and allow one to reproduce (Geary 2005).

These dynamics set the stage for the within-species arms race, whereby cognitively and socially sophisticated individuals are able to outmaneuver and manipulate other individuals in order to gain control of local resources and to exert disproportionate influence over the lives of other people. Much of the competition that might arise under these conditions is muted in modern democracies (Pinker 2011), but even so the relation between social status and health and well-being remains (Marmot 2004). Prior to the rise of these societies, resource control and social influence had substantive effects on survival and reproductive outcomes (Betzig 1986; Hed 1987; Malthus 1798; United Nations 1985), and thus the supporting sociocognitive competencies and brain systems would have necessarily evolved. The gist is that heightened social competition, following the evolution of ecological dominance, accommodates most of the core features of the climatic and ecological theories of human brain evolution (Alexander 1989; Flinn et al. 2005; Geary 2005), as well as other evolutionary models that focus primarily on the importance of social competition during human evolution (e.g., Dunbar 1998; Humphrey 1976).

Fluid Intelligence

Across the various evolutionary theories, the basic definitions of intelligence include planning, foresight, scenario building, and so forth, but seminal psychological research on intelligence is almost never mentioned. One goal of an earlier work was to attempt to meld the modular view of the human mind associated with these various evolutionary models with the more than 100 years of psychometric and more recent cognitive neuroscience research on intelligence (Geary 2005). *The key hypothesis is that fluid intelligence evolved because of the advantages provided by the ability to mentally represent potential changes in social conditions and to engage in explicit problem solving to devise behavioral strategies to cope with these changes.* The evolution of fluid intelligence probably accelerated after hominids achieved ecological dominance (e.g., through tools, fire, shelter) and the resulting uptick in the intensity of social competition (i.e., within-species arms race). I summarize this framework in a later section (see Geary 2005 for full description). In the following, I provide a brief review of empirical studies of general intelligence.

Psychometric Intelligence

In Spearman's (1904) seminal study, children, adolescents, and adults were administered sensory and perceptual tasks and were rated by teachers and peers on their in-school intelligence and out-of-school common sense. Exam scores were also available for the adolescents. Correlations revealed that above-average performance on one task was associated with above-average performance on all other tasks, on exam scores, and for ratings of intelligence and common sense, leading Spearman (1904, p. 285) to conclude "that all branches of intellectual activity have in common one fundamental function (or group of functions)." Spearman termed the fundamental function or group of functions general intelligence or *g*.

Decades later, Cattell and Horn (Cattell 1963; Horn 1968) proposed that general intelligence should be subdivided into *crystallized intelligence* (Gc; learned knowledge and skills) and *fluid intelligence* (Gf). The latter represents a biologically based ability to acquire skills and knowledge. As Cattell (1963, p. 3) stated: "Fluid general ability … shows more in tests requiring adaptation to new situations, where crystallized skills are of no particular advantage." Cattell's description of fluid intelligence is consistent with the gist of all of the various models of hominid brain and cognitive evolution; specifically, this evolution was driven by selection pressures that required humans and our ancestors to cope with variation and novelty in their day-to-day lives. In other words, the ability to anticipate and cope with novelty and change that are central to models of human brain and cognitive evolution is reliably assessed by tests of Gf.

Cognitive Components of Intelligence

Hundreds of studies have focused on identifying the cognitive components of intelligence (Deary 2000; Jensen 1998). These efforts led to the identification of speed of information processing and working memory as core components of intelligence, especially Gf; many of these studies include measures of Gc or those that tap a combination of Gc and Gf and thus are termed "intelligence."

Speed of Processing

The first of three important findings to emerge is that faster speed of cognitive processing is related to higher scores on measures of intelligence (Jensen 1982; Jensen and Munro 1979), although strength of the relation is moderate ($rs \sim 0.3$–0.4). The second is that individuals who are consistently fast in executing the same processes time after time have the higher intelligence scores ($rs \sim 0.4$; Deary 2000; Jensen 1998). Finally, the speed with which individuals can identify very briefly (e.g., 50 ms) presented information (e.g., whether ">" is pointed left or right) is moderately correlated with intelligence (Deary and Stough 1996).

The gist is that performance on psychometric test of intelligence and especially Gf is related to the speed and accuracy with which information is identified and then processed by sensory and perceptual systems. For individuals who score highly on measures of Gf, the processing of this information occurs more rapidly than for other individuals. For all individuals, the information is first implicitly (i.e., below conscious awareness) represented in short-term memory. For the information to become available to conscious awareness and thus amendable to explicit problem solving, it must become represented in working memory.

Working Memory and Problem Solving

The processing of the majority of information gathered by our sensory and perceptual systems and represented in short-term memory occurs automatically and implicitly. Typically, mental or behavioral responses also occur automatically, that is, without the need to engage attentional and working-memory resources (Gigerenzer et al. 1999). These automatic responses are the result of evolved heuristics – fast and efficient behavioral or cognitive responses that require minimal, explicit cognitive resources – or heuristics learned during the life span. However, when information cannot be automatically processed by evolved systems or through access to overlearned information, the result is an automatic shift in attention to this information (Botvinick et al. 2001). Situations that trigger attentional shifts are, by this definition, novel or rapidly changing.

Attentional focus results in an explicit representation of this information in working memory and simultaneous inhibition of irrelevant information (Cowan 1995; Engle et al. 1995). Once active in working memory and available to conscious awareness, the information is amendable to problem solving. The attentional system that controls the explicit manipulation of information during problem solving is called the central executive, which operates on information in several representational systems, including auditory, visual, spatial, or episodic (Baddeley 1986, 2000). Episodic memory binds information from multiple systems and is important for recall of memories of personal experiences (Tulving 2002).

Individual differences in performance on measures of Gf are moderately ($rs \sim 0.5$; Ackerman et al. 2005) to strongly associated with individual differences in working-memory capacity ($rs > 0.8$; Conway et al. 2002). On the basis of these patterns, Horn (1988) and other scientists (Carpenter et al. 1990; Stanovich 1999) have argued that measures of strategic problem solving and abstract reasoning define Gf, and the primary cognitive system underlying problem solving, reasoning, and thus Gf is attention-driven working memory; a proposal regarding the potential relation between speed of processing and working memory is presented in Integration.

Summary Intelligent individuals identify, process, and bind together bits of social and ecological information more easily and quickly than do other people. Their perceptual systems process this information such that it is activated quickly and accurately in short-term memory. If evolved or learned heuristics are available for responding to the situation, then intelligent people will be able to execute these responses more quickly and consistently than other people. If evolved or learned heuristics are not available, there is an automatic shifting of attention to the information represented in short-term memory.

Once attention is focused, intelligent people are able represent more information in working memory than are other people and have an enhanced ability to consciously manipulate this information. The manipulation is guided and constrained by reasoning and inference-making mechanisms (Embretson 1995; Stanovich 1999).

Brain Systems and Intelligence

I have proposed that this attention-driven ability to explicitly represent and systematically and logically manipulate information in working memory are core features of an evolved human ability to adapt to social and ecological variation and novelty within the life span (Geary 2005). If correct, then the evolutionary emergence of Gf should track the above-described evolutionary changes in brain size and EQ. Unfortunately, a direct test of this hypothesis is not possible due to the absence of Gf data in the paleontological record. There are some interesting patterns, nevertheless. I noted earlier that there have been disproportionate increases in the size and organization of some prefrontal and parietal regions, regions that have also been implicated in brain imaging studies of Gf and self-awareness.

Brain Size and Regional Activation

There is a modest relation between gross brain volume and intelligence ($r \sim 0.3$–0.4; Deary 2000; McDaniel 2005; Rushton and Ankney 1996). Recent reviews also indicate modest correlations between the size of specific regions in the prefrontal, parietal, and temporal cortices and potentially the corpus callosum (Deary et al. 2010; Luders et al. 2009; Jung and Haier 2007). Of particular importance may be the size and integration of regions in the prefrontal cortex (areas 6, 9, 10, 32, and 45–47, Fig. 8.2), the attentional control and imagery integration areas of the parietal cortex (areas 7, 39, 40, Fig. 8.2), and the white matter tracts that connect them, which are components of Jung and Haier's parieto-frontal integration theory (P-FIT) of intelligence. Results from functional MRI and other imaging techniques are also consistent with the P-FIT theory; that is, these frontal and parietal regions are typically active while people are engaged in difficult problem-solving and reasoning tasks (e.g., Duncan et al. 2000; Gray et al. 2003; Haier et al. 1988), although other task-specific regions may be engaged as well.

Many of the regions identified as being critical to Gf overlap with those engaged in working-memory tasks, consistent with the previously described correlations, although there appear to be differences as well. In a unique and large-scale study, Barbey et al. (2014) studied deficits in intelligence and the maintenance aspect of working memory (keeping something in mind while engaged in other mental processes) in individuals with focal brain damage. They confirmed the importance of the parieto-frontal network for intelligence, especially in the right hemisphere. The working-memory task was dependent on some of these same regions but also many regions of the left hemisphere (e.g., areas 41, 42, Fig. 8.2) that were not important for intelligence. These results need to be interpreted with some caution, however, because the working-memory task used in this study does not assess all components of working memory (e.g., inhibitory control), and lesion studies may not capture the dynamic integration of areas during complex reasoning and problem solving, as noted in another large-scale lesion study (Gläscher et al. 2009).

In any case, the anterior cingulate (area 32, Fig. 8.2) is included in the P-FIT model and is important because this brain region is engaged when goal achievement requires dealing with some degree of novelty, conflict, or making a difficult decision (Miller and Cohen 2001; Ranganath and Rainer 2003) – these are situations in which a goal cannot be readily achieved by means of heuristics. Areas of the anterior cingulate cortex are thus the potential mechanism that results in the automatic attentional shift to novel, conflicted, or changing information represented in short-term memory and a corresponding activation of the dorsolateral and other prefrontal areas (Botvinick et al. 2001). These areas in turn enable the attentional focus and explicit, controlled problem solving needed to cope with novel situations, resolve conflicts, and make decisions that involve cost-benefit trade-offs (Kerns et al. 2004; Miller and Cohen 2001).

Integration

Brain imaging studies suggest that an integrated system of brain regions supports the explicit controlled problem solving that is the core of fluid intelligence and that many, but not all, of these same regions support working memory (Duncan et al. 2000; Gray et al. 2003; Kane and Engle 2002). High scores on measures of fluid intelligence are associated with activation of the dorsolateral prefrontal cortex and several other brain regions associated with attentional control, including the anterior cingulate cortex and regions of the parietal cortex (Jung and Haier 2007). Many of these same regions contribute to the ability to inhibit irrelevant information from intruding into working memory and conscious awareness (Esposito et al. 1999) and inhibit the execution of evolved or learned heuristics (Geary 2005). Awareness of information represented in working memory and the ability to mentally manipulate this information may result from a synchronization of the prefrontal brain regions that subserve the central executive and the brain regions that process the specific forms of information (e.g., voice, face, object; Damasio 1989; Posner 1994).

An attention-driven synchronization of the activity of the P-FIT network and activation of domain-specific brain regions would be facilitated by faster speed of processing and rich interconnections among these brain regions. The latter are associated with larger brain size and especially a greater volume and myelination of axons (i.e., white matter; Deary et al. 2010). Speed of processing may be important for the synchronization process, because faster speed of processing would enable more accurate adjustments in regional synchronization per feedback cycle. With repeated synchronized activity, the result appears to be the formation of a neural network that automatically links the processing of these information patterns. In other words, speed of processing and an attention-driven working-memory system are not competing explanations of Gf but rather may be coevolved and complementary mechanisms.

Integrated Model: The Motivation to Control

The theoretical model that places general fluid intelligence in the context of the human brain and cognitive evolution can be organized around a "motivation to control." This does not mean individuals necessarily have a conscious, Machiavellian motivation to control others. Rather, the "motivation to control" is a heuristic for conceptualizing the function of evolved traits.

General Theory

The basic proposal is that the brain and mind of all species evolved to process the forms of information (e.g., facial expressions, movement patterns of predators) that covaried with survival and reproductive prospects during the species' evolutionary history. These systems operate implicitly and bias the organism to behave in ways that result in attempts to gain control of these outcomes, such as capturing prey or avoiding being captured by a predator (Gigerenzer et al. 1999; Simon 1956). The gist is consistent with the well-replicated finding that people's subjective well-being and physical health are associated with having some level of control over relationships, events, and resources that are of significance in their life (Heckhausen and Schulz 1995). As noted earlier, in conditions in which we evolved, achieving control of social relationships (e.g., as related to social status) and biological (e.g., food) and physical (e.g., safe shelter, water) resources often meant the difference between living and dying.

The control-related behavioral focus is represented by the apex and adjoining section of Fig. 8.3. The bottom of the figure represents the folk modules that result in implicit and automatic, bottom-up processing of social (e.g.,

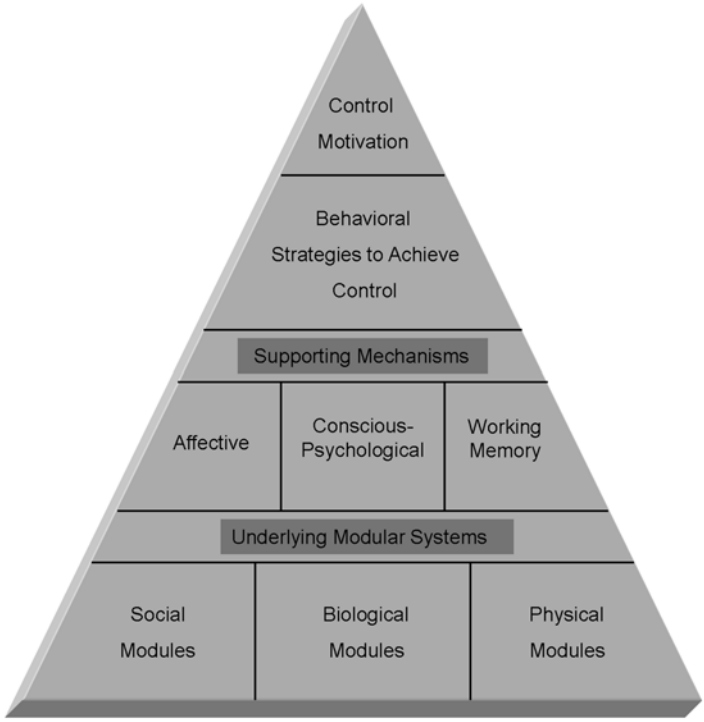


Fig. 8.3 The apex and following section represent the focus of behavior on achieving control of the social, biological, and physical resources that have tended to covary survival and reproductive outcomes during human evolution. The midsection shows the supporting affective, conscious-psychological (e.g., self-awareness), and cognitive (e.g., working memory) mechanisms that support the motivation to control and operate on the modular systems shown at the base (Adapted from Geary 2005, p. 74. Copyright 2005 by American Psychological Association)

facial expressions), biological (e.g., features of hunted species), and physical (e.g., objects potentially useable as tools) information patterns that have tended to be the same across generations and within lifetimes and have covaried with survival or reproductive prospects during human evolution; a corresponding taxonomy of folk abilities is provided elsewhere (Geary 2005; Geary and Huffman 2002). If the evolutionary expansion of the human brain and presumably enhancement of fluid intelligence was driven by the need to deal with rapid change in climatic, ecological, and especially social conditions, then cognitive systems that function to mentally represent and manipulate dynamic change in information patterns should be identifiable. These systems are represented by the center section of Fig. 8.3 and are the key to understanding the evolution of fluid intelligence.

Conscious-Psychological

The core of the conscious-psychological mechanism is the human ability to form a conscious, explicit mental representation of situations that are centered on the self and one's relationships with other people or one's access to biological and physical resources. The representations often involve a form of mental time travel; that is, mental simulations of past, present, or potential future states and can be cast as visual images, in language, or as episodic memories (Suddendorf and Corballis 1997; Tulving 2002). The key is the ability to create a mental representation of a desired or fantasized state and to compare this to a mental representation of one's current situation. These are self-centered conscious-psychological representations of present and potential future states that are of personal significance and are the content on which more conscious and effortful reasoning and problem-solving processes are applied (Evans 2002; Stanovich and West 2000), with the goal of reducing the difference between the current and desired states (Geary 2005).

I have argued elsewhere that the evolution of self-awareness and infusion of this awareness into these mental simulations is best explained by social selection pressures, that is, awareness of oneself is particularly important when attempting to cope with other people with conflicting interests who know your personality, behavioral biases, etc. (Geary 2005). Mental simulation of the behavior of other people and especially of the self engages an identifiable brain network that includes the prefrontal (area 32, 10, Fig. 8.2) and parietal (area 7) cortices (Gallagher and Frith 2003; Lou et al. 2004). These systems appear to reconstitute the activity of the brain regions that were engaged during personal experiences or activate more abstracted representations of common features of these experiences (Damasio 1989) and are highly integrated with the prefrontal regions described in the next section (Miller and Cohen 2001).

Cognitive Mechanisms

The cognitive mechanisms include working memory, attentional control, and the ability to inhibit automatic processing of evolved or learned behavioral responses or cognitive biases (e.g., attributional biases; Baddeley 2000; Cowan 1995). The mechanisms also include the ability to systematically problem-solve and to reason about patterns represented in working memory (Newell and Simon 1972). In effect, working memory and attentional and inhibitory control are the content-free mechanisms that enable the integration of a current conscious-psychological state with memory representations of related past experiences and the generation of mental models or simulations of potential future states (Alexander 1989).

Attentional and executive control is dependent on several regions of the prefrontal cortex, such as the dorsolateral regions, as well as areas of the parietal cortex (Kane and Engle 2002). These areas are highly integrated with other regions of the prefrontal cortex, such as the ventromedial areas (e.g., area 11) of both hemispheres, the right frontal pole (area 10), and areas of the parietal cortex (e.g., area 7) that support social cognition, including a sense of self (Tulving 2002). The integration of these areas is consistent with a coevolutionary process, whereby working memory and fluid intelligence are readily activated during the processing of social information and self-relevant situations. My proposal is that these systems are coactivated in social contexts that require novel behaviors and in circumstances when one is attempting to outwit competitors or avoid being outwitted by them (Geary 2005).

Affective Mechanisms

Affective mechanisms include emotions, that is, observable behaviors (e.g., facial expressions), and feelings – conscious representations of an emotional state or other conditions that can potentially influence the individual's well-being (Damasio 2003). Emotions result in observable feedback to others, and feelings provide unobservable feedback to the individual (Campos et al. 1989). The latter is a useful indicator of the effectiveness of control-related behavioral strategies and an indicator of the potential benefits of a simulated behavior. Positive feelings provide reinforcement when strategies are resulting in a reduction in the difference between the current and desired state, and negative feelings promote disengagement when behaviors are not resulting in this end (Gray 1987). The supporting brain systems, such as the amygdala (not shown in Fig. 8.2), are predicted to function in part to amplify attention to evolutionarily significant forms of information and produce emotions, feelings, and corresponding behavioral biases that are likely to automatically reproduce outcomes that have covaried with survival or reproduction during hominid evolution.

Conclusion

Since the emergence of australopithecines more than 4 MYA, the hominid brain has tripled in size, with much of the change occurring after the emergence of *Homo* and especially during the past several hundred thousand years. The proposed selection pressures driving this change include climatic variation (Potts 1998; Vrba 1995), the complexities of hunting and foraging (Kaplan et al. 2000), and social dynamics (Alexander 1989; Dunbar 1998; Humphrey 1976). It may be that each class of pressure contributed to the evolution of the human brain over the past 4 million years, but the most interesting issue concerns the particularly rapid and substantial increases in brain volume and EQ following the emergence of *H. erectus* and continuing through to modern humans (Bailey and Geary 2009; Pearce et al. 2013). Alexander's ecological dominance model provides a way of integrating these selection pressures and for making inferences about when the various pressures may have been most critical in driving human brain and cognitive evolution. On the basis of this model, my colleagues and I have argued that the complexity and dynamics of social competition and cooperation within and between groups is likely to have been the most potent selection pressure for human brain and cognitive evolution since the emergence of *H. erectus* (Flinn et al. 2005; Geary 2005).

The combination of selection pressures and especially those related to social competition can be integrated with psychological research using the motivation to control framework. The folk physical systems at the base of Fig. 8.3 and described elsewhere support the evolution of modularized brain and cognitive systems that enable humans to conceptualize, construct, and use very complicated tools (Geary 2005), consistent with Kaplan et al.'s (2000) hunting hypothesis and Alexander's (1989) ecological dominance model. As noted, the model also predicts the evolution of modularized folk psychological or sociocognitive competencies that are uniquely human, which would include language, theory of mind, face processing, among others (Geary 2005). Much of human behavior and cognition and perhaps all of that of most other species can be accommodated by the fast, implicit functioning of such modularized systems, for example, resulting in rapid responses to predators or threatening facial expressions.

The motivation-to-control model includes modules in these domains as well as mechanisms that enable organisms to anticipate and cope with variation, novelty, and change within their life span. The critical differences comparing humans to other species appear to be in the abilities to generate representations of the self in working memory – self-awareness – and to mentally time-travel (Suddendorf and Corballis 1997; Tulving 2002). The combination enables the generation of self-centered mental models, that is, a conscious-psychological (explicit) representation of past, present, or potential future situations that are of personal relevance. Although climatic and ecological conditions can create variation and novelty, the most dynamic and variable conditions faced by humans are those that arise from the competing interests of other people. My proposal is that the evolved function of these mental models is to generate a self-centered simulation of the "perfect" world, one in which other people behave in ways consistent with one's best interest, and biological and physical resources are under one's control. The function of mental simulations is to create and rehearse strategies that can be used to reduce the difference between this perfect world and current conditions. The cognitive systems that evolved to support the use of these self-centered mental models are known as working memory and general fluid intelligence.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60.

Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: The brain and digestive system in human and primate evolution. *Current Anthropology, 36*, 199–221.

Alexander, R. D. (1989). Evolution of the human psyche. In P. Mellars & C. Stringer (Eds.), *The human revolution: Behavioural and biological perspectives on the origins of modern humans* (pp. 455–513). Princeton, NJ: Princeton University Press.

Alroy, J. (2001). A multispecies overkill simulation of the end-Pleistocene megafaunal mass extinction. *Science, 292*, 1893–1896.

Asfaw, B., White, T., Lovejoy, O., Latimer, B., Simpson, S., & Suwa, G. (1999). Australopithecus garhi: A new species of early hominid from Ethiopia. *Science, 284*, 629–635.

Ash, J., & Gallup, G. G., Jr. (2007). Paleoclimatic variation and brain expansion during human evolution. *Human Nature, 18*, 109–124.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.

Bailey, D. H., & Geary, D. C. (2009). Hominid brain evolution: Testing climatic, ecological, and social competition models. *Human Nature, 20*, 67–79.

Barbey, A. K., Colom, R., Paul, E. J., & Grafman, J. (2014). Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Structure and Function*, *219*, 485–494.

Barton, R. A. (1996). Neocortex size and behavioural ecology in primates. *Proceedings of the Royal Society of London B, 263*, 173–177.

Betzig, L. L. (1986). *Despotism and differential reproduction: A Darwinian view of history*. New York: Aldine.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. [Comparative localization of the cerebral cortex based on cell composition.] Leipzig: Barth.

Campos, J. J., Campos, R. G., & Barrett, K. C. (1989). Emergent themes in the study of emotional development and emotion regulation. *Developmental Psychology, 25*, 394–402.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of processing in the Raven Progressive Matrices Test. *Psychological Review, 97*, 404–431.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1–22.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163–183.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33*, 25–62.

Damasio, A. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*. Orlando: Harcourt.

Darwin, C. (1859). *The origin of species by means of natural selection*. London: John Murray.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Darwin, C., & Wallace, A. (1858). On the tendency of species to form varieties, and on the perpetuation of varieties and species by natural means of selection. *Journal of the Linnean Society of London, Zoology, 3*, 45–62.

Deary, I. J. (2000). *Looking down on human intelligence: From psychophysics to the brain*. Oxford: Oxford University Press.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist, 51*, 599–608.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience, 11*, 201–211.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology, 6*, 178–190.

Duncan, J., Rüdiger, J. S., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F. H., & Emslie, H. (2000). A neural basis for general intelligence. *Science, 289*, 457–460.

Elton, S., Bishop, L. C., & Wood, B. (2001). Comparative context of Plio-Pleistocene hominin brain evolution. *Journal of Human Evolution, 41*, 1–27.

Embretson, S. E. (1995). The role of working memory capacity and general control processes in intelligence. *Intelligence, 20*, 169–189.

Engle, R. W., Conway, A. R. A., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. *Psychological Science, 6*, 122–125.

Esposito, G., Kirkby, B. S., van Horn, J. D., Ellmore, T. M., & Berman, K. F. (1999). Context-dependent, neural system-specific neurophysiological concomitants of ageing: Mapping PET correlates during cognitive activation. *Brain, 122*, 963–979.

Evans, J. S. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin, 128*, 978–996.

Falk, D., Redmond, J. C., Jr., Guyer, J., Conroy, G. C., Recheis, W., Weber, G. W., & Seidler, H. (2000). Early hominid brain evolution: A new look at old endocasts. *Journal of Human Evolution, 38*, 695–717.

Flinn, M. V., Geary, D. C., & Ward, C. V. (2005). Ecological dominance, social competition, and coalitionary arms races: Why humans evolved extraordinary intelligence. *Evolution and Human Behavior, 26*, 10–46.

Foley, R. (1987). Hominid species and stone-tool assemblages: How are they related. *Antiquity, 61*, 380–392.

Foley, R., & Lahr, M. M. (1997). Mode 3 technologies and the evolution of modern humans. *Cambridge Archaeology Journal, 7*, 3–36.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences, 7*, 77–83.

Geary, D. C. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.

Geary, D. C. (2010). *Male, female: The evolution of human sex differences* (2nd ed.). Washington, DC: American Psychological Association.

Geary, D. C., & Huffman, K. J. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psychological Bulletin, 128*, 667–698.

Gigerenzer, G., Todd, P. M., & ABC Research Group (Eds.). (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gläscher, J., Tranel, D., Paul, L. K., Rudrauf, D., Rorden, C., Hornaday, A., & Adolphs, R. (2009). Lesion mapping of cognitive abilities linked to intelligence. *Neuron, 61*, 681–691.

Gray, J. A. (1987). Perspectives on anxiety and impulsivity: A commentary. *Journal of Research in Personality, 21*, 493–509.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316–322.

Haier, R. J., Siegel, B. V., Jr., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., Browning, H. L., & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied using positron emission tomography. *Intelligence, 12*, 199–217.

Heckhausen, J., & Schulz, R. (1995). A life-span theory of control. *Psychological Review, 102*, 284–304.

Hed, H. M. E. (1987). Trends in opportunity for natural selection in the Swedish population during the period 1650–1980. *Human Biology, 59*, 785–797.

Hill, K., Boesch, C., Goodall, J., Pusey, A., Williams, J., & Wrangham, R. (2001). Mortality rates among wild chimpanzees. *Journal of Human Evolution, 40*, 437–450.

Holloway, R. L. (1973). Endocranial volumes of early African hominids, and the role of the brain in human mosaic evolution. *Journal of Human Evolution, 2*, 449–459.

Holloway, R. (1996). Evolution of the human brain. In A. Lock & C. R. Peters (Eds.), *Handbook of human symbolic evolution* (pp. 74–116). New York: Oxford University Press.

Holloway, R. L., & de al Coste-Lareymondie, M. C. (1982). Brain endocast asymmetry in pongids and hominids: Some preliminary findings on the paleontology of cerebral dominance. *American Journal of Physical Anthropology, 58*, 101–110.

Holloway, R. L., & Kimbel, W. H. (1986). Endocast morphology of Hadar hominid AL 162–28. *Nature, 321*, 536–537.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review, 75*, 242–259.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 645–685). New York: Plenum.

Hsieh, C.-h., Reiss, C. S., Hunter, J. R., Beddington, J. R., May, R. M., & Sugihara, G. (2006). Fishing elevates variability in the abundance of exploited species. *Nature, 443*, 859–862.

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). New York: Cambridge University Press.

Huxley, T. H. (1863). *Evidence as to man's place in nature*. New York: Appleton and company.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). New York: Springer.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.

Jensen, A. R., & Munro, E. (1979). Reaction time, movement time, and intelligence. *Intelligence, 3*, 121–126.

Jerison, H. J. (1973). *Evolution of the brain and intelligence.* New York: Academic.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging data. *Behavioral and Brain Sciences, 30*, 135–187.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*, 637–671.

Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology, 9*, 156–185.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*, 1023–1026.

Levine, B. (1999). Self-regulation and autonoetic consciousness. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn conference* (pp. 200–214). Philadelphia: Psychology Press.

Lou, H. C., Luber, B., Crupain, M., Keenan, J. P., Nowak, M., Kjaer, T. W., Sackeim, H. A., & Lisanby, S. H. (2004). Parietal cortex and representation of the mental self. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 6827–6832.

Luders, E., Narr, K. L., Thompson, P. M., & Toga, A. W. (2009). Neuroanatomical correlates of intelligence. *Intelligence, 37*, 156–163.

Malthus, T. R. (1798). *An essay on the principle of population as it affects the future improvement of society with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers*. London: Printed for J. Johnson, in St. Paul's church-yard.

Marmot, M. (2004). *The status syndrome: How social standing affects our health and longevity*. New York: Henry Holt and Company, LLC.

Martin, P. S. (1973). The discovery of America: The first Americans may have swept the Western Hemisphere and decimated its fauna in 1000 years. *Science, 179*, 969–974.

McDaniel, M. A. (2005). Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence, 33*, 337–346.

McHenry, H. M. (1994). Tempo and mode in human evolution. *Proceedings of the National Academy of Sciences USA, 91*, 6780–6786.

Miller, E. K., & Cohen, J. D. (2001). An integration of theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Mithen, S. (1996). *The prehistory of the mind: The cognitive origins of art and science*. New York: Thames and Hudson.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Pearce, E., Stringer, C., & Dunbar, R. I. M. (2013). New insights into differences in brain organization between

Neanderthals and anatomically modern humans. *Proceedings of the Royal Society B, 280*, 2130–2140.
Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. New York: Viking.
Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences USA, 91*, 7398–7403.
Potts, R. (1998). Variability selection in hominid evolution. *Evolutionary Anthropology, 7*, 81–96.
Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience, 4*, 193–202.
Rushton, J. P., & Ankney, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin & Review, 3*, 21–36.
Semaw, S., Rogers, M. J., Quade, J., Renne, P. R., Butler, R. F., Dominquez-Rodrigo, M., Stout, D., Hart, W. S., Pickering, T., & Simpson, S. W. (2003). 2.6-million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *Journal of Human Evolution, 45*, 169–177.
Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.
Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology, 15*, 201–293.
Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645–726.
Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs, 123*, 133–167.
Tobias, P. V. (1987). The brain of *Homo habilis*: A new level of organization in cerebral evolution. *Journal of Human Evolution, 16*, 741–761.
Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25.
United Nations. (1985). *Socio-economic differentials in child mortality in developing countries*. New York: Author.
Vrba, E. S. (1995). The fossil record of African antelopes (Mammalia, Bovidae) in relation to human evolution and paleoclimate. In E. S. Vrba, G. H. Denton, T. C. Partridge, & L. H. Burckle (Eds.), *Paleoclimate and evolution, with emphasis on human origins* (pp. 385–424). New Haven, CT: Yale University Press.
Wallace, A. R. (1911). *The world of life: A manifestation of creative power, direction mind and ultimate purpose*. New York: Moffat, Yard & Company.
Wood, B., & Collard, M. (1999). The human genus. *Science, 284*, 65–71.
Wrangham, R. W., Holland Jones, J., Laden, G., Pilbeam, D., & Conklin-Brittain, N. (1999). The raw and stolen: Cooking and the ecology of human origins. *Current Anthropology, 40*, 567–594.

Part III

Theories of Intelligence

9 Intelligence Defined: Wundt, James, Cattell, Thorndike, Goddard, and Yerkes

John D. Greenwood

Most texts in the history of psychology credit Francis Galton (1822–1911) and Alfred Binet (1857–1911) as the first to develop theories of intelligence as well as instruments for its measurement. However, credit should probably go to the Victorian polymath Herbert Spencer (1820–1903) as the first to develop a substantive theory of intelligence (Guilford 1967) and one which implied—at least for Spencer—individual, racial, and species differences in intelligence. Spencer treated the empiricist principle of association by contiguity as the foundation of intelligence in both animals and humans:

> Hence the growth of intelligence at large depends upon the law, that when any two psychical states occur in immediate succession, an effect is produced such that if the first subsequently recurs there is a certain tendency for the second to follow it. (1855, p. 530)

Spencer held that intelligence is determined by the quantity and quality of adaptive associations made by organisms to their environment, by the "the continuous adjustment of internal relations to external relations" (1855, p. 374), which he believed was in turn determined by neurophysiological complexity. Consequently, he maintained that intelligence is a function of brain size.

Like Charles Darwin (1809–1882), Spencer believed in strong psychological continuity between humans and other animals: that the psychological capacities of humans and other animals differ in degree but not in fundament kind.[1] While for Darwin this just meant that complex and distinctively human capacities such as reasoning and language could be attributed to higher animals in at least insipient form, Spencer held that humans and other animals differ only in the complexity of their associative processes. For Spencer, the complex capacities of humans such as reasoning and language are merely elaborations of the basic associative processes common to all animals.

Spencer's theory was enormously influential in its day, prompting theorists such as John Hughlings Jackson (1835–1911) and Ivan Sechenov (1829–1910) to develop reflexive sensorimotor theories of the nervous system, which treated cognitive capacities such as reasoning and language as merely the pinnacles of a hierarchy of increasingly more complex and sophisticated reflexive machinery (Jackson

[1] "The difference in mind between man and the higher animals, great as it is, is certainly one of degree and not of kind. We have seen that the senses and intuitions, the various emotions and faculties, such as love, memory, attention, curiosity, imitation, reason, &c., of which man boasts, may be found in an incipient, or even sometimes in a well-developed condition, in the lower animals." (Darwin 1871, p. 105)

J.D. Greenwood, Ph.D. (✉)
CUNY Graduate Center, 365 Fifth Avenue, New York, NY 10016, USA
e-mail: jgreenwood@gc.cuny.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_9,

1931; Sechenov 1863). They developed theories of strong continuity between so-called "higher" cognitive processes such as reason and language and "lower" associative or reflexive processes, mirroring Spencer and Darwin's theory of strong continuity between human and animal psychology.[2]

Thus in a very real sense, scientific psychological theories of intelligence in the nineteenth century began with theories of animal intelligence, and the question of the relation between human and animal intelligence remained a live issue for many early scientific psychologists.

Wundt and James: Apperception and Similarity

Wilhelm Wundt (1832–1920) and William James (1842–1910) are generally held to be the founding fathers of scientific psychology in Germany and America (respectively). However, neither followed Spencer in treating intelligence in terms of the complexity of association by contiguity. This is because both Wundt and James rejected the principles of strong psychological continuity—between humans and animals and between cognitive and associative processes—championed by evolutionary theorists like Spencer and experimental physiologists like Hughlings Jackson and Sechenov. Both Wundt and James championed the autonomy of a scientific psychology devoted to the study of conscious experience by rejecting what they perceived to be reductive challenges to the reality and efficacy of human conscious experience.

Thus Wundt, for example, maintained that the creative and selective attentional processes of apperception, which shape human perception and cognition, are distinct from the forms of association that humans share with animals. Like Conwy Lloyd Morgan (1852–1936), Wundt held that all animal psychology and behavior could "be accounted for by the simple laws of association" (Wundt 1863, p. 350) and treated apperception as the highest stage of distinctively human psychological evolution. Consequently, Wundt had little to say about intelligence per se, except to distinguish the forms of apperceptive intelligence—involving the creative synthesis of relational psychological elements—distinctive of humans from the forms of associative intelligence characteristic of both humans and animals.

William James managed to persuade the President and Board of Trustees at Harvard that the "new psychology" provided an intellectual bastion against the materialist threats of evolutionary psychology and experimental physiology, which according to his future colleague at Yale, George Trumbull Ladd (1842–1921), threatened to reduce conscious experience to "a stream of mechanically associated 'epiphenomena,' thrown off by the molecular machinery of the cerebral hemispheres" (1895, p. x). Like Wundt and Lloyd Morgan, James held that all animal psychology and behavior could be explained in terms of association by contiguity and claimed that distinctive human cognition went beyond this to the recognition of similarities and analogies:

> We may then, we think, consider it proven that *the most elementary single difference between the human mind and that of brutes lies in this deficiency on the brute's part to associate ideas by similarity*—characters, the abstraction of which depends on this association, must in the brute always remained drowned, swamped in the total phenomenon which they help constitute. (James 1890, p. 360)

While James did not develop a theory of intelligence per se, it is clear that he thought that the most intelligent men and women were those who were able to most fully develop their ability to form associations by similarity:

> Genius then, as has been already said, is identical with the possession of similar association to an extreme degree. (1890, p. 360)

[2]Although it is doubtful if Darwin himself thought that rationality and language are just more complex elaborations of associative or reflexive processes. That is, he seems to have held that cognition and association are psychologically discontinuous, even though he held that the psychology of humans and animals is strongly continuous: he seems to have supposed that the *discrete* higher cognitive capacities of humans—such as rationality and language—could be found in animals, in at least incipient form.

Although neither Wundt nor James developed substantive theories or measures of intelligence, a number of their students did and made major contributions to debates about the nature of intelligence in the early twentieth century. To these we now turn.

Cattell and Mental Testing

James McKeen Cattell (1860–1944) was one of Wundt's earliest Leipzig students and his first full-time American student. He also became Wundt's first research assistant, based upon Cattell's own recommendation (Boring 1957). While at Leipzig, Cattell did experimental studies of the time taken to complete "cerebral operations," a topic suited to Wundt's program of research on mental chronometry, based upon measures of reaction time.[3] Cattell's paper "On the Time It Takes to See and Name Objects" was published in Wundt's journal *Philosophical Studies* in 1885 (Cattell 1885); a shorter version appeared in the British journal *Mind* a year later (Cattell 1886).

Most unusually, Wundt also allowed Cattell to pursue research on individual differences in reaction time, a topic that was anathema to Wundt, who refused to allow students like Edward B. Titchener (1867–1927) and Lightner Witmer (1867–1956) to participate in reaction time experiments because they were not properly "calibrated" (trained) to reproduce the supposedly universal measures of sensory and motor reaction times (O'Donnell 1985). This may have been because Wundt respected Cattell's intellectual independence and self-confidence, qualities Wundt characterized as "typically American."

There was, however, one major theorist who recognized individual differences in sensory and motor reaction times, and that was Francis Galton (1822–1911), with whom Cattell went to study in England after he completed his degree in Leipzig in 1886. Galton, who was Darwin's half cousin, maintained that individual differences in sensorimotor reaction times were a natural consequence of the chance distribution of inherited characteristics described by Darwin in *On the Origin of Species* (Darwin 1857) and set about measuring these differences in the general population. Galton gathered data on nearly 20,000 people at his anthropometric laboratories at the International Health Exhibition in London (in 1884) and the Science Galleries of the South Kensington Museum (in 1888), employing physical and sensory acuity measures, such as head size, physical strength, visual and auditory acuity, and reaction time.

Galton claimed that sensory acuity was correlated with intelligence and thus could serve as a convenient indirect measure of intelligence. Galton also later claimed that he had demonstrated the correlation, based upon the statistical measures of correlation that he developed toward the end of the 1880s (Galton 1888), although he seems to have come to this conclusion sometime earlier. For example, in *Inquiries Into Human Faculty* (1883), Galton announced:

> The trials I have as yet made on the sensitivity of different persons confirm the reasonable expectation that it would on the whole be highest among the intellectually ablest. (1883, p. 20)

In this peculiar fashion, the measures of sensory acuity and reaction time developed by Wundt and his students in the Leipzig laboratory were appropriated by Galton, and later by Cattell, as measures of intelligence.

Galton also employed his newly developed statistical measures—and his pioneering use of twin studies—to supposedly demonstrate that human intelligence is largely determined by heredity. While the statistical calculations of *Natural Inheritance* (Galton, 1889) are impressive, the data on which they were founded are doubtful, based upon family records anonymously submitted by correspondents hoping to win the cash prizes for best entries offered by Galton (Boakes 1984).

Given his commitment to the hereditarian determination of intelligence, Galton was

[3]This program was based upon the complication experiments of the Dutch physiologist Franciscus Cornelis Donders (1818–1889), in which the time taken for components of a complex mental task was calculated by subtracting the time taken for other components of the task.

consequently dismissive of optimistic utilitarian theories—such as those advanced by John Stuart Mill (1806–1873)—that held that all human beings are capable of attaining the same intellectual and moral levels, given similar nurturing experiences. As Galton put it (once again in advance of his statistical calculations):

> I have no patience with the hypothesis …that babies are born pretty much alike, and that the sole agencies in creating differences between boy and boy, and man and man, are steady application and moral effort. It is in the most unqualified manner that I object to pretensions of natural equality. (Galton 1869, p. 12)

Although Galton did not believe that the intelligence levels of men and women could be significantly raised through education, he did think they could be raised through selective breeding. Like others who reflected on Darwin's theory of evolution, Galton recognized that natural selection operating on chance variations would not ensure the evolution of socially desirable characteristics like high intelligence and moral virtue and indeed might very well lead to their attenuation, if "idiots and imbeciles" were allowed to overbreed. Galton coined the term eugenics (Greek for "wellborn") to describe his recommended form of artificial selection, designed to produce more intelligent and productive human stock (as farmers used selective breeding to produce desired qualities in their animal stock, such as high quality of fleece in sheep and body mass in cattle).

Originally, Galton recommended a form of positive eugenics. He thought that those persons identified as the most intelligent (via his sensory acuity measures) should be encouraged to breed, and to breed regularly, via financial inducements provided by the government. However, in reaction to the moral panic created by the failure of the mighty British Empire to crush a nation of farmers in the Second Boer War (1899–1902), Galton and his protégé Karl Pearson (1857–1936) recommended a form of negative eugenics through institutionalization and sterilization of the "idiots and imbeciles."

Cattell was greatly impressed by Galton's work, which reinforced his own interest in individual differences. He also embraced Galton's theory of the hereditarian determination of intelligence and contributed his small share to the program of positive eugenics by offering each of his seven children one thousand dollars if they married college professors. When he returned to America for a position at the University of Pennsylvania, Cattell began to use the techniques he had developed at Leipzig for his own anthropometric studies and initiated a program of "mental testing" based upon psychophysical measures of grip strength, speed of movement, skin sensitivity, and sensory and motor reaction times. In a paper published in *Mind* in 1890 (Cattell 1890), Cattell extolled the utility of such "mental tests," and when he moved to Columbia in 1891, he subjected hundreds of students to them.

Like Galton, Cattell assumed that his tests of sensory acuity were indicators of intelligence and consequently assumed that there would be a significant correlation between test scores and academic performance. However, at the turn of the century, one of Cattell's own students, Clark Wissler (1870–1947), tried to demonstrate this by measuring the degree of correlation between performance on 21 of Cattell's mental tests and course grades, employing the Galtonian measures of correlation (including Pearson's newly developed correlation coefficient). To Cattell's consternation, Wissler (1901) found no correlation between the test scores and the grades, nor between the tests scores themselves (although he did find correlations between the student's grades). This effectively brought to an end to the program of the Galtonian anthropometric intelligence testing in the United States.[4] However, intelligence testing would shortly resurface in an entirely different guise.

Spearman and the "g" Factor

Wissler's studies were not quite the end of the Galtonian psychometric story, however, although they are commonly supposed to have been.

[4]Although it was continued for a few years by Joseph Jastrow (1863–1944) at the University of Michigan.

Charles Spearman (1863–1945), after serving as a regular officer in the British Army for 15 years, resigned his commission in 1897 to study for a doctoral degree in Wundt's Laboratory. While at Leipzig, Spearman was surprised to learn about Wissler's results, but eventually came to the conclusion that Wissler's data were unreliable and that he had underestimated the degree of relationship between them. Using a corrective statistical formula, Spearman was able to demonstrate positive correlations between the test scores and between the test scores and academic grades. He parleyed this success into his theory that the correlation between various mental tests could be explained by reference to a single, unitary capacity, or general factor "g," which he identified via a new statistical technique that came to be known as factor analysis.[5]

Spearman called his theory, which he also developed and published while a student at Leipzig, as the "law of the universal unity of the intellectual function":

> Whenever branches of intellectual activity are at all dissimilar, then their correlations with one another appear wholly due to their being all variously saturated with some common fundamental Function (or group of Functions). (Spearman, 1904, p. 124)

Like Galton and Cattell, Spearman also believed that this general factor ''g'' is innate:

> G is in the normal course of events determined innately; a person can no more be trained to have it in higher degree than he can be trained to be taller. (Spearman 1931, in Deary et al 2008, p 157)

After Leipzig, Spearman returned to England and took up a position at University College, London, where he continued to develop his theory of intelligence. His theory was later criticized by Sir Geoffrey Thompson (1881–1955), who argued that although g was a statistical reality, it did not designate a unitary intelligence factor, but a variety of highly correlated intellectual skills (Thomson 1916). Spearman did not always insist on the unitary nature of the g factor, at least in his later work (Spearman 1925). He believed g to be grounded in two "ubiquitously" cooperating abilities: an educative ability to make "meaning from confusion" and a reproductive ability to recall that meaning. And although Spearman remained a champion of intelligence testing and of the view that the innate g factor accounted for most individual differences of intelligence, he thought that intelligence testing had no place in schools, which ought to be engaged in maximizing the varied native abilities of individual students.[6]

Thorndike and Connectionism

Edward L. Thorndike (1874–1949) was a student of William James and received his master's degree at Harvard in 1897. Like James (and Wundt), he was not impressed by anecdotal reports of the apparently intelligent behavior of animals, such as the apparent rationality of animals reported by Darwin in *The Descent of Man, and Selection in Relation to Sex* (Darwin 1971)[7] or their apparent knowledge of mathematical and mechanical principles reported by George Romanes (1848–1894) in *Animal Intelligence* (Romanes 1882).[8] Thorndike disparaged anecdotal accounts of animal behavior because he maintained they were generally unrepresentative of animals' cognitive abilities:

> Dogs get lost hundreds of times and no one ever notices it or sends an account of it to a scientific magazine. But let one find his way from Brooklyn to Yonkers and the fact immediately becomes a circulating anecdote. Thousands of cats on thousands

[5]Strictly speaking, Spearman developed a two-factor theory of intelligence, postulating a general factor "g" common to all tasks requiring intelligence and a factor "s" specific to different kinds of intellectual tasks (Sternberg 2003).

[6]Unlike his later colleague at University College, Sir Cyril Burt (1883–1971), who promoted the 1940s British government program of intelligence testing in schools—the "11 plus" exam—on the basis of which students were streamed into academic or trade classes.

[7]Darwin claimed that most people would agree that "animals possess some power of reasoning," on the grounds that they "may constantly be seen to pause, deliberate, and resolve" (1871, p. 46).

[8]Darwin bequeathed his notebooks on animal behavior to Romanes after the two men became close friends toward the end of Darwin's life.

> of occasions sit helplessly yowling, and no one takes thought of it or writes to his friend, the professor; but let one cat claw at the knob of a door supposedly as a signal to be let out, and straightway this cat becomes the representative of the cat-mind in all the books…In short, the anecdotes give really the *abnormal* or *supernormal* psychology of animals. (Thorndike 1911, pp. 23–25)

Thorndike insisted that the scientific study of animal "intelligence" should be based upon carefully controlled experimental studies, which he began with chickens in William James's basement and continued at Columbia University when Cattell offered him a fellowship there in 1897. At Columbia, Thorndike began a series of experiments in which food-deprived cats learned to escape from specially constructed slatted cages or "puzzle boxes" to gain a food reward, described in his doctoral dissertation of 1898, and published later that year as *Animal Intelligence: An Experimental Study of the Associative Processes in Animals* (Thorndike 1898) as a monograph supplement in *Psychological Review*[9].

Thorndike's studies were based upon Lloyd Morgan's (1894) explanation of how his dog Toby had learned to lift the latch on the gate of the back courtyard of his house to escape into the street. Based upon his repeated observation of the dog's behavior, Lloyd Morgan dismissed Romanes's (1882) explanation of such behavior in terms of the dog's understanding of mechanical principles and noted how after accidentally stumbling on a means of lifting the latch by a movement of its head the dog had managed to learn after repeated trial and error how to open the gate. Similarly, Thorndike's cats initially clawed at the bars and pushed their paws between them, until they accidentally hit on the movement required to release the latch. Like Morgan, Thorndike found that the animals took progressively less time to hit on the required behavior over a series of trials, until eventually they produced the learned behavior the moment they were placed in the box.

On the basis of these experiments, Thorndike articulated what he called the "law of effect," linking learning with reinforcement, and the "law of exercise," linking learning with repetition.

Thorndike was at pains to insist that learning was not based upon imitation or insight and that connections between behavior and response were "stamped in" by reinforcement and repetition. Consequently, he called his theory "connectionism" and on the basis of his limited experiments maintained that all animal and human behavior could be explained in terms of the laws of effect and exercise and instinct:

> The higher animals, including man, manifest no behavior beyond expectations from the laws of instinct, exercise and effect. (Thorndike 1911, p. 274)

While he was later forced to modify the law of effect and drop the law of exercise, Thorndike continued to insist that learning was based upon connections or associations. So although he rejected Darwin and Romanes's attribution of higher cognitive processes such as rationality and mechanical understanding to animals, he maintained—with Spencer—that higher cognitive processes are nothing more than complex forms of connection or association, based ultimately upon the same psychological and physiological principles:

> …the higher forms of intellectual operations are identical with mere association or connection forming, depending upon the same sort of physiological connections but requiring *many more of them*. (Thorndike et al. 1926, p. 415)

Consequently, he supposed that an individual's level of intelligence was determined by the number of connections that individual was capable of making:

> By the same argument the person whose intellect is greater or higher or better than that of another person differs from him in the last analysis in having, not a new sort of physiological process, but simply a larger number of connections of the ordinary sort. (Thorndike et al. 1926, p. 415)

While he recognized the difficulty of determining the nature and number of connections that needed to be made to in order to execute a complex human cognitive process such as language comprehension, Thorndike developed his own intelligence tests when he took up a full-time

[9] It was republished as an independent monograph in 1911 (Thorndike 1911).

position at Teachers College, where he remained for the rest of his career, devoting much of his time to educational psychology and the development of tests for education (mainly on reading and writing). Although he came to his theoretical conception of human intelligence by a different route than Galton, Cattell, and Spearman, Thorndike was equally committed to the view that intelligence was largely determined by hereditary, publishing *Heredity, Correlation and Sex Differences in School Abilities* in 1903 (Thorndike 1903).

Goddard and the Binet-Simon Intelligence Scale

A good case can be made for the claim that Granville Stanley Hall (1844–1924), rather than William James, deserves to be credited as the founding father of scientific psychology in America. Although Hall was himself a student of James, he was the first to complete a PhD at Harvard on a psychological topic (with a dissertation on spatial perception). He was also the first to set up a fully developed psychological laboratory and PhD program in psychology at Johns Hopkins University in 1884,[10] which he transferred to Clark University in 1888. Hall is mainly remembered for his contributions to educational psychology and (like Wundt and James) had little to say about intelligence and its measurement, but he did have two graduate students who made significant contributions to the field.

Henry Goddard (1866–1957) graduated from Clark University in 1899 and became Director of the Research Laboratory for the Study of Feeblemindedness at the Vineland Training School for Feeble-Minded Boys and Girls in New Jersey in 1906. He translated the 1908 Binet-Simon scale for measuring children's intelligence into English and also the revised 1911 Binet-Simon scale. This became the standard measure of intelligence in the United States until Lewis Terman (1877–1956), a Clark University graduate of 1905 who secured a position as Professor of Child Study at Stanford University, brought out what came to be known as the Stanford-Binet scale in 1916.

In 1905 Alfred Binet (1857–1911) and his research assistant Theodore Simon (1873–1961) had developed a test for measuring children's intelligence, as a means of assessing the child's "mental level" (the tests were designed to ensure that most children at any age would test at the appropriate mental level). Unlike Galton and Cattell's tests, Binet and Simon's tests were based on direct measures of intellectual abilities such as comprehension, problem solving, and logical and analogical reasoning—all forms of intelligent judgment. Their scales consisted of thirty items, ranked in order of difficulty, so that everyone could do the easier ones, and children scored progressively better with every increase in "mental level."

Binet and Simon's test items represented what Spearmen called a "hodgepodge" of factors, but they did the job for which they were designed, namely, to provide an objective means of identifying children in need of remedial education. Unlike Spearman, Binet and Simon did not suppose that they were measuring a unitary capacity, far less one that is innately determined. On the contrary, they designed programs of remedial education for those children who tested below the mental level of their age group, designed to increase their mental level through special training methods called "mental orthopedics" (which involved exercises in attention, will, and discipline). The Binet-Simon scale was administered to French schoolchildren between 1905 and 1908 and was judged to be a great success in Europe. It was seen as an objective measure of intelligence that was easy to administer and was translated into many languages.

Binet and Simon were careful to stress the limitations of their scales, given their belief in the malleability of intelligence and the inherent margin of error. However their cautious approach was discarded when Goddard and Terman

[10] Although William James set up a laboratory of sorts at Harvard in 1875, it was little more than a small collection of instruments housed in a stairwell closet (Hall 1923, p. 218), and Harvard did not attain an independent (of philosophy) psychology department until 1937.

brought the Binet-Simon scales to the United States—both followed Galton, Cattell, and Spearman in supposing that intelligence was a unitary ability, which was largely determined by heredity. In 1914, William Stern (1871–1938) had introduced the theoretical notion of a mental quotient, defined as a child's mental age—as determined by their performance on the Binet-Simon scale—divided by their chronological age (Stern 1914), and Terman defined the intelligence quotient (IQ) as Stern's mental age multiplied by 100, so that the average intelligence quotient for any mental age would be 100. Goddard and Terman used these IQ scales to define levels of intelligence and "feeblemindedness," or as Terman put it bluntly in his PhD dissertation at Clark University, "genius and stupidity" (Terman 1906).

Goddard was convinced of the utility of the tests as measures of intelligence and trained teachers at local schools to administer them. He was not surprised to find that the scores for inmates at Vineland were much lower than those of children in the regular public schools but was concerned to discover that many of the public schoolchildren scored lower than their age norms, which Goddard took as a disturbing indicator of the extent of feeblemindedness in the general population.

Goddard also impressed the immigration officers at Ellis Island with his apparent ability to identify feebleminded immigrants by sight and then have his identification "objectively" confirmed by their low scores on the Binet-Simon tests. Goddard trained assistants to identify the feebleminded and administer the tests, and eventually they came to be employed by the inspectors themselves, although some objected that they included questions (e.g., about the New York Giants) that they would not have been able to answer when they first came to the country. Nevertheless, the use of psychological testing methods for screening immigrants was judged to be a great success, and the number of persons deported on grounds of feeblemindedness increased dramatically. Again, Goddard drew a pessimistic conclusion about the high percentage of feeblemindedness among recent immigrants of "poor stock," that is, among the increasing number of immigrants from Eastern and Southern Europe, as opposed to the earlier and supposedly "superior stock" from Northern and Western Europe.

Goddard had also read the works of Galton and Mendel and had noted that many brothers and sisters of the children at Vineland were also feebleminded themselves, which he operationally defined as having an IQ of less than 70. In 1911, he set out to demonstrate the inheritance of feeblemindedness, which he believed to be caused by a recessive gene, by exploring the family tree of a 22-year-old Vineland girl whom Goddard had named Deborah Kallikak. She performed on the Binet-Simon scale at the mental level of 9 years, on the basis of which Goddard classified her as a moron, a technical term (meaning dull) introduced by Goddard to describe those with an IQ of between 50 and 70.

Goddard set about exploring her family tree, which he found could be traced back to one Martin Kallikak, a Revolutionary War soldier. Martin had married an upstanding Quaker girl, who had borne him seven children. The descendants of this "good" side of the family tree had gone on to become upright citizens, such as lawyers, doctors, judges, teachers, and landed gentry. However, Martin Kallikak had also dallied with a serving wench of loose morals, who had borne him an illegitimate son, Martin Kallikak junior, who fathered ten children. The descendants of this "bad" side of the family turned out to be horse thieves, brothel owners, prostitutes, and alcoholics. Goddard claimed to have discovered a high incidence of feeblemindedness among the offspring of the serving wench and a low incidence among the offspring of the upstanding Quaker girl.

This study, published in 1912 as *The Kallikak Family: A Study in Feeblemindedness* (Goddard 1912), confirmed Goddard in his belief that feeblemindedness and intelligence are inherited, and that feeblemindedness was the root cause of licentiousness and criminality. Although Goddard promoted the study as a "natural experiment," there was no attempt to control for environmental or social differences, and Goddard and his coworkers, like many asylum and prison superintendents of his day, simply equated immoral and criminal behavior with feeblemindedness. Nevertheless, Goddard's conclusions were widely accepted and regularly cited, and his study spawned a spate of similar studies supposedly demonstrating the link between feeblemindedness and social degeneration (Zenderland 1998). Many also accepted the implied threats to the "national stock" posed by the overbreeding of what Galton had called "the idle and the infirm" and the influx of recent immigrants of poor stock.

For Goddard, the moral and social implications were as clear as they were for Galton, namely the need for eugenic programs designed to prevent the breeding of the feebleminded. Goddard served as psychology representative on the 1911 Committee to Study and to Report on the Best Practical Means of Cutting Off the Defective Germ-Plasm in the American Population, established by the Eugenics Section of the American Breeders Association, which recommended the segregation and sterilization of mental defectives, mercifully drawing the line against euthanasia. Goddard also served on the 1913 Committee for the Heredity of Feeblemindedness, which included fellow psychologists Edward Thorndike, Lewis Terman, and Robert M. Yerkes (1876–1956), as well as the inventor Alexander Graham Bell (1847–1922) and the Harvard physiologist Walter B. Canon (1871–1945), which also recommended that "defective classes" be eliminated through sterilization.

Such eugenic ideas were not themselves a product of intelligence testing, but had been in circulation since the beginning of the century. Charles Davenport (1866–1944), the author of *Eugenics: The Science of Human Improvement by Better Breeding* (1911), had founded the Eugenics Records Office at Cold Spring Harbor, New York, in 1910, and similar societies and organizations had been founded in Canada and Europe. However, the major impetus of the development of eugenics legislation in the United States was the result of the program of intelligence testing in the United States Army conducted during the First World War, engineered by the psychologist Robert M. Yerkes.

Yerkes and the Army Testing Project

Robert M. Yerkes originally came to Harvard as a graduate student in zoology, but the philosopher Josiah Royce (1855–1916) persuaded him to combine his interests in zoology and psychology to study comparative psychology. Yerkes transferred to the philosophy department, then headed by James's successor, Hugo Münsterburg (1863–1916), and received his PhD degree in 1902. Münsterburg managed to secure Yerkes an assistantship and later an assistant professorship in the department, where he developed "criteria of the psychic" (Yerkes 1905), which supposedly provided researchers with objective grounds for the attribution of higher cognitive states to animals, as a means of assessing their intelligence.[11] However, Münsterburg objected to Yerkes's failure to follow his advice to devote more of his time to educational psychology and threatened to shut down the comparative psychology program. Fortunately, Yerkes secured a position as state psychologist at Boston Psychopathic Hospital just before the First World War, which also enabled him to keep a half-time teaching position at Harvard at double his former pay.

In later years, Yerkes became known for his seminal contributions to comparative psychology, but during his time at Boston Psychopathic Hospital, he developed his own intelligence test, the Yerkes-Bridges Point Scale of Intelligence (Yerkes et al. 1915). Yerkes was also a born organizer and administrator, and his great chance came in April 1917, when Titchener's Experimentalists were meeting at Harvard. Yerkes was serving that year as President of the American Psychological Association and, on receiving the news that America had entered the

[11] John B. Watson (1878–1958), who was a friend of Yerkes, was so skeptical of his criteria for the objective determination of animal cognition and intelligence that he began to promote the claim that animal psychology should be restricted to the description of observable stimulus–response sequences (Watson 1909), a position he extended to human psychology 4 years later in his "Psychology as a Behaviorist Views It" lecture at Columbia (Watson 1913).

First World War, chaired a special session on the proposed contribution of psychologists to the war effort. Shortly afterward, Yerkes traveled to Canada to study the contribution of Canadian psychologists to the war effort, and later that month, at the meeting of the APA Council in Philadelphia, he formed a committee to explore ways in which American psychologists could contribute. This committee, whose members included Cattell, Hall, Thorndike, and Watson, suggested the development of psychological tests to facilitate the selection of officers and the discharge of feebleminded recruits. Yerkes promptly formed the Committee on Methods of Psychological Examining for Recruits, whose members included Goddard and Terman, which spent 2 weeks at Vineland creating group intelligence tests for the army and running trials at local institutions and army bases.

When the army finally approved what became known as the Army Testing Project, some 400 commissioned psychologists administered group intelligence tests to some two million soldiers between 1917 and 1919: the Alpha[12] test to literate soldiers and the Beta (pictorial) test to illiterates. The mass testing was of doubtful military utility, and the army discontinued the project at the end of the war (Samelson 1977), although it did serve to promote the public perception of the utility of intelligence testing and the professional status of psychologists—as Cattell put it, the Army Testing Project "put psychology on the map" (Cattell 1922, p. 5). Yerkes was appointed to the National Research Council, where he worked on the development of the National Intelligence Test (National Research Council 1920), administered to over seven million schoolchildren in the 1920s.

However, the most important outcome of the Army Testing Project was the alarming finding that around half the army recruits tested at or below the level of moron. In his final report on the project, Yerkes concluded that "feeblemindedness …is of much greater frequency than has been previously supposed" (Yerkes 1921, p. 789). This led to a moral panic analogous to that generated in Britain by the poor performance of the British army during the Second Boer War and promoted similar fears about the decline in "national efficiency." These fears were stoked by books like Madison Grant's *The Passing of the Great Race* (Grant 1916), Goddard's *Human Efficiency and Levels of Intelligence* (Goddard 1920), and Carl Brigham's *A Study of American Intelligence* (Brigham 1923, with a foreword by Yerkes), which claimed that the average intelligence of Americans had declined since the turn of the century due to the overbreeding of recent immigrants from Eastern and Southern Europe, which was itself seen as a sign of feeblemindedness.

While there were those who urged caution in drawing conclusions from the army data (Boring 1923a) and those who disputed Goddard's (Goddard 1919) claim that they demonstrated that the average intelligence of adult Americans was below the level of moron (Freeman 1922; Lippmann 1922), they were drowned out by more strident calls for immigration quotas and programs of sterilization of the feebleminded.

In 1924, the Congress passed the National Origins Act, which imposed quotas on the nationality of immigrants based upon the 1890 census (i.e., before the wave of Eastern and Southern European immigration at the turn of the century). Harry Laughlin (1880–1943) of the Eugenics Records Office had testified before the Congressional Immigration and Naturalization Committee earlier that year, claiming that the American gene pool was being "polluted" by the growing numbers of intellectually inferior immigrants.

By the end of the 1920s, close to 30 states[13] had passed laws legitimizing the compulsory sterilization of the feebleminded, a procedure

[12]Carl Brigham (1890–1943), who worked on the Army Alpha test, later adapted it as an admissions test when he joined the faculty at Princeton University, where it became known as the Scholastic Admissions Test (SAT) and later as the Scholastic Assessment Test.

[13]The first state law licensing the compulsory sterilization of confirmed criminals, idiots, imbeciles, and rapists was passed by Indiana in 1907, on the recommendation of Harry Clay Sharp (1869–1940), a pioneer of vasectomy, who originally used the procedure to treat masturbation, which he believed to be a major cause of intellectual degeneracy. The 1907 Indiana Law was struck down by

inflicted on at least 12,000 inmates by 1930. In 1926, the US Supreme Court put its seal of approval on the eugenic use of the procedure by sanctioning the sterilization of Carrie Buck (1906–1983), a 17-year-old girl who was an inmate of the Virginia Colony for the Epileptic and Feebleminded. Carrie and her mother, who was also an inmate, were judged to be feeble-minded, as was Carrie's illegitimate daughter Vivian, and Carrie was sterilized by court order under Virginia's Eugenical Sterilization Act.[14] The case was appealed to the US Supreme Court, which upheld the Virginia Law, with Oliver Wendell Holmes Jr. (1841–1935), who wrote the majority opinion in the case of Buck vs. Bell in 1927, famously declaring:

> It is better for all the world if instead of waiting to execute degenerate offspring for crime, or let them starve for their imbecility, society can prevent those who are manifestly unfit from continuing their kind. Three generations of imbeciles are enough. (Buck vs. Bell 1927, p. 207)

Eventually, the eugenic excesses of the Nazis—who did endorse euthanasia on a grand scale—demonstrated the dangers of this over-reaching science and dampened the enthusiasm of many former supporters. Many original psychological advocates, such as Goddard and Brigham, recanted their original positions.[15] And Yerkes, although he continued to believe in genetically determined racial differences in intelligence, recognized that the topic had become too hot to handle and shifted his attention back to comparative psychology. Nonetheless, the quota system established by the National Origins Act remained in place until 1965, and the sterilization laws remained on the books for many years later, with the State of Virginia repealing the last law as late as 1981.

Conclusion: So What Is Intelligence?

Beyond the *Sturm und Drang* over the inheritance of intelligence and feeblemindedness, what was learned about the nature of intelligence itself? Not much it seemed. In a 1921 study published in the *Journal of Educational Psychology* that canvassed the definitions of intelligence by 14 "experts" (Thorndike et al 1921, there was remarkably little agreement, and many practitioners of intelligence testing rested content with Boring's (1923b) operational definition of intelligence as what intelligence tests measure, which seemed sufficient to guarantee its reference while saying nothing about its nature. Nevertheless, one can discern at least the outline of a common theme running through the works of Wundt, James, Spearman, Thorndike, and the developers of the Binet-Simon tests of intelligence. Wundt's apperceptive synthesis of relational elements, James' discernment of similarities, Spearman's "education" of relations and correlates, Thorndike's formation of connections, and Binet and Simon's trio of "comprehension, invention, and direction" all point to some form of cognitive achievement involving the discernment or determination of connections and configurations. However, it would take the development of cognitive psychology in the later part of the century to begin to specify the relevant cognitive processes, while still leaving open and contentious the original question of whether "higher" cognitive processes such as logical reasoning and abstract thought are continuous with "lower" forms of association and whether human intelligence is continuous with animal intelligence (Mitchell et al. 2009; Penn et al. 2008; Shanks 2007).

the state supreme court in 1921, but was quickly replaced by another that survived legislative challenge.

[14]There is little evidence that Carrie was feebleminded, and she seems to have been institutionalized on grounds of sexual promiscuity—commonly treated as an indicator of feeblemindedness—on the basis of her pregnancy, despite the fact that this was the result of a rape by a nephew of the foster parents who committed her. Nor is there much evidence that Carrie's daughter Vivien was feebleminded. Although she died at the age of eight of an intestinal disease, her first grade report card indicated that she was a solid B student, with an A in deportment, who once made the honor role.

[15]Brigham also repudiated the use of the Scholastic Achievement Test as a measure of intelligence (Angier et al. 1926) and opposed its use as the basis of a National Educational Testing Service (Brigham 1938).

References

Angier, R. P., MacPhail, A. H., Rogers, D. C., Stone, C. J., & Brigham, C. C. (1926). *Scholastic aptitude tests: A manual for the use of schools*. New York: College Entrance Examination Board.

Boakes, R. (1984). *From Darwin to behaviorism: Psychology and the minds of animals*. Cambridge, MA: Cambridge University Press.

Boring, E. G. (1923a). Facts and fantasies of immigration. *New Republic, 36* 245–246.

Boring, E. C. (1923b). Intelligence as the tests test it. *New Republic, 36*, 35–37.

Boring, E. G. (1957). *A history of experimental psychology* (2nd ed.). New York: Appleton Century Crofts.

Brigham, C. C. (1923). *A study of American intelligence*. Princeton: Princeton University Press.

Brigham, C. C. (1938). Letter to J. B. Conant, 3 January. Princeton: ETS Archive:Brigham file. Buck v. Bell – 274 U.S. 200 (1927).

Cattell, J. M. (1885). Über die Zeit der Erkennung und Benennung von Scriftzeichen, Bilden und Farben. *Philosophische Studien, 2*, 635–650.

Cattell, J. M. (1886). On the time it takes to see and name objects. *Mind, 11*, 63–65.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15*, 373–381.

Cattell, J. M. (1922). The first year of the psychological corporation. Report at annual meeting of the psychological corporation, December 1. Lewis M. Terman Papers, Stanford University Archives.

Darwin, C. (1859). *On the origin of species by means of natural selection*. London: John Murray.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Davenport, C. (1911). *Eugenics: The science of human improvement by better breeding*. New York: Henry Holt & Co. Deary

Deary I. J., Lawn, M., & Bartholomew, D. J. (2008). A conversation between Charles Spearman, Godfrey Thomson, and Edward L. Thorndike: The International Examinations Inquiry Meetings 1931–1938. *History of Psychology, 11*, 122–142.

Freeman, F. N. (1922). The mental age of adults. *Journal of Educational Research, 6*, 441–444.

Galton, F. (1869). *Hereditary genius*. London: Macmillan.

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Galton, F. (1888). Co-relations and their measurement: Chiefly from anthropometric data. *Proceedings of the Royal Society, 45*, 135–145.

Galton, F. (1889). *Natural inheritance*. London: Macmillan.

Goddard, H. H. (1912). *The Kallikak family: a study in the heredity of feeblemindedness*. New York: Macmillan.

Goddard, H. H. (1919). *Psychology of the normal and subnormal*. New York: Dodd, Mead.

Goddard, H. H. (1920). *Human efficiency and levels of intelligence*. Princeton: Princeton University Press.

Grant, M. (1916). *The passing of the great race*. New York: Scribner's.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Hall, G. S. (1923). *Life and confessions of a psychologist*. New York: Appleton.

Intelligence and Its Measurement: A Symposium. (1921). *Journal of Educational Psychology*, 12, 123–147, 195–216, 271–275.

Jackson, J. H. (1931). *Selected writings of Hughlings-Jackson* (Ed., J. Taylor, Vols. 1–2). London: Hodder & Stroughton.

James, W. (1890). *The principles of psychology* (Vol. 1–2). New York: Holt.

Ladd, G. T. (1895). *Philosophy of mind: An essay on the metaphysics of psychology*. New York: Scribner's.

Lippmann, W. (1922). The mental age of Americans. *New Republic, 32*, 213–215.

Mitchell, C. J., Houwer, J. D., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–246.

Morgan, C. L. (1894). *Introduction to comparative psychology*. London: Walter Scott.

National Research Council. (1920). *Fourth annual report*. Washington, DC: Government Printing Office.

O'Donnell, J. M. (1985). *The origins of behaviorism: American psychology, 1870–1920*. New York: New York University Press.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and non-human minds. *Behavioral and Brain Sciences, 31*, 109–178.

Romanes, G. J. (1882). *Animal intelligence*. New York: Appleton.

Samelson, F. (1977). World War I intelligence testing and the development of psychology. *Journal of the History of the Behavioral Sciences, 13*, 274–282.

Sechenov, I. (1863). *Reflexes of the brain* (ed.: G. Gibbons, Trans: S. Belsky). Cambridge, MA: MIT Press, 1965.

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 37B*, 1–21.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology, 15*, 201–292.

Spearman, C. (1925). *The abilities of man: Their nature and measurement*. London: McMillan.

Spencer, H. (1855). *Principles of psychology*. London: Longmans.

Stern, W. (1914). The psychological methods of testing intelligence (trans: G. M. Whipple). Baltimore: Warwick and York.

Sternberg, R. J. (2003). Intelligence. In D. K. Freedheim & I. B. Weiner (Eds.), *Handbook of psychology* (History of psychology, Vol. 1). Hoboken: Wiley.

Terman, L. M. (1906). Genius and stupidity. *Pedagogical Seminary, 13*, 307–373.

Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology, 8*, 271–281.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in

animals. *Psychological Review, Monograph Supplement, 2*, (8), 1–109.
Thorndike, E. L. (1903). *Heredity, correlation and sex differences in school abilities*. New York: Macmillan.
Thorndike, E. L. (1911). *Animal intelligence*. New York: Hafner.
Thorndike, E. L., Henmon, V. A. C., & Buckingham, B. R. (eds.) (1921). Intelligence and its measurement: A Symposium. Baltimore.
Thorndike, E. L., Bregman, E. D., Cobb, M. V., & Woodyard, E. I. (1926). *The measurement of intelligence*. New York: Teachers College Press.
Watson, J. B. (1909). A point of view in comparative psychology. *Psychological Bulletin, 6*, 57–58.
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158–177.
Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review Monograph Supplements, 3*(6).
Wundt, W. (1863) *Lectures on human and animal psychology* (trans: J. E. Crighton & E. B. Titchener). New York: Macmillan, 1894.
Yerkes, R. M. (1905). Animal psychology and the criterion of the psychic. *Journal of Philosophy, Psychology and Scientific Methods, 2*, 527–528.
Yerkes, R. M. (Ed.) (1921). Psychological examining in the United States Army. *Memoirs of the National Academy of Sciences, 15*, 1–890.
Yerkes, R. M. (1923). *Introduction to C. C. Brigham's A study of American intelligence*. Princeton: Princeton University Press.
Yerkes, R. M., Bridges, J. W., & Hardwick, R. S. (1915). *A point scale for measuring mental ability*. Baltimore: Warwick & York.
Zenderland, L. (1998). *Measuring minds: Henry Herbert Goddard and the origins of American intelligence testing*. Cambridge, MA: Cambridge University Press.

10 Piaget's Theory of Intelligence

Ulrich Müller, Kayla Ten Eycke, and Lesley Baker

> *[I]ntelligence, the most plastic and at the same time most durable structural equilibrium of behavior, is essentially a system of living and acting operations*
>
> (Piaget 1976a, p. 7)

Readers may be puzzled by the quote above because the way in which Piaget uses the term intelligence is rather broad and does not correspond with the more tightly defined way in which it is used today. Puzzlement likely will make way to outright confusion when the reader attempts to make sense of Piaget's definition of intelligence: "Intelligence constitutes the state of equilibrium towards which tend all the successive adaptions of a sensori-motor and cognitive nature, as well as all assimilatory and accommodatory interactions between the organism and the environment" (Piaget 1976a, p. 11). This quote is taken from the book *The Psychology of Intelligence*, but despite its title, the following pages of the book might do little to lift the reader's confusion. In this book, there is no description of any method by means of which intelligence might be assessed in order to assign individuals a particular number (their IQ), nor does it refer to individual differences in intelligence, nor does it identify specific cognitive processes such as working memory or processing speed as being involved in intelligence, nor is there a discussion of a specific neural basis of intelligence, nor is any attempt being made to understand intelligence in the context of the selection pressures exerted by the environment of evolutionary adaptedness. Instead, in the first part of the book the reader encounters a classification of theories of intelligence, a discussion of old theories of intelligence (e.g., Russell's theory of logical atomism, the Wùrzburg school of thought psychology, Gestalt psychology), and remarks on the relation between logic and psychology that culminate in a mathematical treatment of higher states of equilibrium.

The goal of this chapter is to clarify Piaget's theory of intelligence. We attempt to do this by first showing how Piaget's conception of intelligence builds on that of his contemporaries. Next, we describe how his conception of intelligence is contextualized within his larger theoretical framework, his genetic epistemology. This leads us to a discussion of core features of Piaget's theoretical framework such as self-organization,

U. Müller (✉) • L. Baker
Department of Psychology, P. O. Box 1700 STN CSC, University of Victoria, Victoria, BC V8W 2Y2, Canada
e-mail: umueller@uvic.ca; lesleyb@uvic.ca

K. Ten Eycke
Behavioural Research Unit, Alberta Children's Hospital, University of Calgary, 2888 Shaganapi Trail NW, T3B 6A8 Calgary, AB, Canada
e-mail: k.teneycke@ucalgary.ca

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_10,

assimilation and accommodation, and operative and figurative aspects of intelligence. We next summarize the four major stages of the development of intelligence. Finally, we describe how Piaget's theory of intelligence relates to semiotic function, affectivity, and social interaction.

Piaget's Definition of Intelligence

Similar to his contemporaries (e.g., see Binet 1894, 1975), Piaget uses the term intelligence in a broad sense, a common practice commented on already by the French historian Taine (1872, p. vii): "If I am not mistaken, we mean nowadays by intelligence, what was formerly called Understanding or Intellect—that is to say, the faculty of knowing." Reflecting such a broad sense of intelligence, Piaget (1976a, pp. 8–9) identifies a central feature of intelligence as a movement toward increasing spatiotemporal distances in the functional interaction between subject (i.e., person, animal) and world.

Piaget (1976a, p. 9) notes that a problem arises when we want to draw a line and precisely demarcate behavior that is intelligent from behavior that is not. He illustrates this by comparing a variety of different definitions of intelligence. For example, Karl Bühler (1933) distinguishes between three stages of purposeful behavior: instinct, training, and intellect. Instinctive behaviors are innate, rigid, and executed in the same manner in all members of a species. Training involves trial-and-error learning such that successful behaviors are reinforced and unsuccessful ones are eliminated, which allows individuals to adapt to new situations. True intelligence emerges at the stage of the intellect at which the individual "*makes discoveries by means of insight and reflection*. INVENTION, in the true sense of the term, is the biological achievement of the intellect" (Bühler 1933, p. 10; emphases in original). Piaget (1976a, pp. 9–10) contrasts Bühler's definition with that of Claparède (1917), who argues that the defining feature of intelligence is the adaptation to new situations. As a result, Claparède considers trial-and-error behavior as intelligent behavior and distinguishes it from instinct and habit (see Piaget 1963, pp. 395–407, for an extended discussion of Claparède's theory of intelligence).

Given these contradictory ways of demarcating intelligent and unintelligent behaviors, Piaget (1976a, p. 10) saw himself confronted with the following alternatives:

> [E]ither we must be satisfied with a functional definition at the risk of encompassing almost the entire range of cognitive structures, or else we must choose a particular structure as our criterion, but the choice remains arbitrary and runs the risk of overlooking the continuity which exists in reality.

Piaget (1976a) resolved this issue by combining functional continuity and structural change and defined intelligence in terms of the direction in which development moves, "without insisting on the question of boundaries, which become a matter of stages of successive forms of equilibrium" (p. 10). From the functional perspective, behavior becomes more intelligent as the spatiotemporal distance between person and world increases. At the same time, functional continuity implies that there is neither an absolute starting point for intelligence nor a predetermined end point: "it is an ultimate goal, and its origins are indistinguishable from those of sensori-motor adaptation in general or even from those of biological adaptation itself" (Piaget 1976a, p. 7). From the point of view of structure, earlier stages of development are characterized by a lack of reversibility, with intelligence moving toward increasingly reversible mobility. Let us now examine this definition in the context of Piaget's theoretical framework.

Genetic Epistemology

Piaget (1950a, b, c) called his theoretical framework genetic epistemology. Here, genetic specifies developmental. Epistemology refers to the study of the nature, sources, scope, and validity of knowledge. Usually, epistemology is considered a branch of philosophy, and empirical research is argued to have no bearing on epistemological questions (Hamlyn 1971). Piaget, however, did not believe that epistemological issues fall under the sole jurisdiction of philosophy. One important reason for this is that knowledge itself is in constant flux and always remains incomplete. In this context, Piaget (1972a, p. 2) approvingly quotes the neo-Kantian philosopher Natorp (1910, pp. 14–15):

> Like Kant, we start with the actual existence of knowledge and seek the basis from there. But what is this existence since, as we know, knowledge is constantly evolving? Progression, method is everything… in consequence, the existence of knowledge cannot be comprehended except as a *fieri* (i.e., to be made, to become; authors' note). This *fieri* alone is the fact. Any entity (or object) which knowledge attempts to crystallize must dissolve again in the current of development. It is in the last phase of this development, and in this alone, that we have the right to say: "this is (a fact)." What we can and must seek, then is the law underlying this process.

Hence, if constant evolution is constitutive of scientific knowledge, as witnessed in the natural and human sciences, and even in logic and mathematics (Piaget 1950a, b, 1972a, b), then the study of the conditions of the possibility of knowledge must include the development of knowledge. The study of the development of knowledge, in turn, falls under the purview of the empirical sciences.

Thus, although Piaget is mostly known as a child psychologist, the study of cognitive development in children was for Piaget only a means to address epistemological issues (Vonéche and Vidal 1985). Piaget's focus on epistemological issues explains why he was not at all interested in determining the cognitive level of an individual child or in examining interindividual differences in cognition (see Bringuier 1980, p. 86). Instead, he was interested in what is common to all persons at a specific level of thinking—what Piaget (1987) referred to as the epistemic subject. The epistemic subject is able to attain states of knowledge, and the goal of genetic epistemology is to explain how the attainment of this knowledge is possible. Knowledge, however, "is not to be naively equated with mere belief (or the brute factual existence of a cognitive structure): knowledge has an inescapable normative dimension, one concerning concepts like evidence, objectivity, rationality, validity, truth, etc." (Kitchener 1993, p. 141).

Piaget argued that the normative dimension of knowledge cannot be reduced to causality. He often used the necessary truth of arithmetic judgments as an example to illustrate this irreducibility:

> [the] truth of 2+2=4 is not the 'cause' of the truth of 4–2=2 in the same way that a cannon causes the movement of two billiard balls, or a stimulus is one of the causes of a reaction: the truth … of 2+2=4 'implies' that of 4–2=2, which is quite a different matter (Piaget 1968, p. 187).

The normative dimension of knowledge, then, puts a further constraint on the genetic explanation of knowledge: the explanation must provide for the possibility of the emergence of normative knowledge. Causal, reductionist explanations (e.g., in terms of neurophysiological causality) will not suffice. However, Piaget did not deny that states of knowledge and intelligence have a biological dimension: intelligence has a "dual nature," it is "both biological and logical" (Piaget 1976a, p. 3). Indeed, the major challenge is to explain how intelligence as a system of living and acting operations can be rooted in biology yet incessantly generate novelties and result in rigorous (i.e., logically necessary) and normative knowledge. For Piaget, this challenge can be met only when life is conceived as self-organization.

Self-Organization

It is generally accepted that psychological development has a biological basis. The question is how the relation between biology and psychological development should be conceptualized. Piaget used the concept of self-organization to characterize the relation between biology and psychological development (see Piaget 1971; see Boom 2009; Chapman 1992). At the biological level, self-organization is the process by which a system perpetually reconstitutes its processes (e.g., metabolic cycles) and elements (e.g., cells) in order to preserve its continuous functioning. This self-organizing activity is not something additional or external to the living system; rather, it pervades matter as active form and it arises when the reciprocal interactions between the elements or subsystem of the system lead to the formation of a higher-order system that regulates the elements or subsystems (Piaget 1971, p. 327 fn.). Living systems are self-organizing systems; in exchange processes with their environment, they spontaneously reproduce their organization.

The idea of self-organization can be traced back to Aristotle, and a brief excursion into how he explicated self-organization will help clarify its conceptual status. For Aristotle, form and matter are inseparable; matter is always already organized (Aristotle 1956, 412b, 4–25). Each thing has its substantial form or entelechy; psyche is the entelechy of living beings. Realized structure (entelecheia) of natural kinds must be viewed as the result of immanent processes (their energeia) (Aristotle 1971, 1046a:30, 1050a:22).

Psyche is best interpreted as self-organizing activity that preserves its own organization (Aristotle 1971, 1048b27; see Hübner 1999). Organic self-preservation has its goal inside itself, and this goal is achieved with the occurrence of organic self-preservation. Self-organizing activity preserves its organic body and the disposition that it exercises (Aristotle 1956, 417b2-7; 1971, 1048b18-34). Thus, psyche fulfills the body and preserves the body and the body's disposition to be psyche, and, because psyche is organic self-preservation, body and psyche possess the same being. As a consequence, Aristotle characterizes a living being by the fact that final cause and formal cause are fused together: the goal or telos is the form of the thing; the form of a living being is the continuous self-organizing activity; therefore, the telos of the living being is to preserve its form. A self-organizing system is like a physician healing himself (Aristotle 1970, 199b, 27–32). Machines, by contrast, are not self-organizing systems; their functioning does not result in their reconstitution, and formal cause and final cause are not fused together.

Piaget expanded on Aristotle by putting an evolutionary spin on the idea of self-organization. As he put it, "the very nature of life is constantly to overtake itself" (Piaget 1971, p. 362). Phylogenetically, the interaction between organism and environment leads to the emergence of higher-order self-regulatory processes, which, on the level of cognitive functioning, reflect the basic mechanisms of self-organization and, at the same time, constitute the most complex instruments for regulating the exchange with the environment (Piaget 1971). Moreover, cognitive functioning reflects reason in a double sense: it is the product of reason that is intrinsic to nature (i.e., in the logic of self-organization) and through cognitive functioning reason in nature becomes conscious of itself.

Thus, similar to Binet (Binet and Simon 1916, pp. 141, 153; see Bennour and Vonéche 2009), Piaget considered intelligence a biological adaptation. Self-organizing activity is the biological foundation and origin of intelligence, and cognitive processes are the outcome of and extend the processes of organic self-organization. Cognitive processes extend the processes of organic self-organization by using and adapting to new circumstances the different systems of organic self-regulation that can be found on the genetic, morphogenetic, physiological, and nervous levels. In support of this claim, Piaget (1971) describes many functional and structural analogies between cognitive and organic functioning. Central among these analogies is the triad of assimilation, accommodation, and scheme.

Assimilation, Accommodation, and Scheme

The complementary functions of assimilation and accommodation describe the general characteristics of the exchange between organism and environment (Piaget 1963, 1970, 1971). Assimilation is the aspect of an organism's activity wherein elements of the environment are integrated into the organism's preexisting organizational structures (i.e., the relations between elements). Accommodation, on the other hand, provides the material for the structuring activity of assimilation. Accommodation is the aspect of the activity wherein an organism's existing schemes are differentiated and modified in response to the environment. For example, a preexisting metabolic cycle assimilates particular nutrients by breaking them down into the elements that contribute to the continued functioning of the living system. The assimilatory cycle needs to be modified when the organism encounters a new nutrient (accommodation) (Piaget 1963). Assimilation and accommodation maintain the equilibrium between an organism and its environment.

Assimilation and accommodation at the psychological level extend the physiological interactions between the organism and the environment because their functioning no longer depends on the incorporation of material elements but now incorporates informational content (Piaget 1963, 1971). At the psychological level, schemes are characterized by what is repeatable in actions and thought processes (i.e., internalized actions). Assimilation refers to the incorporation of new information into already existing schemes, a process giving meaning to the content (Piaget 1963, 1985). For example, when a baby grasps a rattle, the rattle is assimilated to the grasping scheme and thereby attains the functional meaning of being "graspable." Assimilation always uses the existing psychological schemes; its functioning carries the history of the subject's interaction with the world into each particular act. For example, an infant who has differentiated various ways of interacting with the rattle will have different action potentialities available compared to an infant who has not.

Accommodation refers to the modification of existing schemes to account for particular features of the object or situation. Because schemes are structures with varying degrees of generality, applying them to particular situations always requires an adjustment or accommodation. Accommodation thus particularizes the general schemes, supplies them with specific content, and modifies them in doing so (e.g., the preexisting grasping scheme needs to be modified, becoming more specific to take into account the particular spatial position of the rattle).

Assimilation, accommodation, and scheme are inseparable. Assimilation is always a structuring activity because it involves integrating content into existing schemes; thus, structures do not exist independently of structuring activity: "Assimilation is hence the very functioning of the system of which organization is the structural aspect" (Piaget 1963, p. 410). At the same time, the incorporation of new elements leads to the modification of the scheme and thus to accommodation. Accommodation brings about adaptation to the environment, but this adaptation is always a function of the structuring activity of assimilation. The act of assimilating objects to schemes is rather complex because it involves affect, sensation, (internalized) motor elements, and perception (Piaget 1963, 1981).

Operative and Figurative Aspects of Intelligence

Closely related to the concepts of assimilation and accommodation are Piaget's (1969) notions of figurative and operative aspects of intelligence. The figurative aspect of intelligence includes the functions of perception, imitation, imagery, and (in part) language that are supplied by the accommodatory aspect of activity (Piaget 1969; Piaget and Inhelder 1971). The figurative aspect provides signifiers, which, in turn, bear on the "states" of reality and provide data on which the structuring activity of assimilation acts. For example, an infant may perceive a rattle and, assimilating it to an action scheme, she recognizes rattle as something that can be shaken (i.e., the sight of the rattle serves as a signifier of what can be done with it).

In contrast, the operative aspect of intelligence refers to the transforming and form-giving, or structuring, aspect of knowledge (Piaget and Inhelder 1971). It includes sensorimotor actions, internalized actions that are carried out mentally, and operations. Operations are internalized actions that have become reversible because they are organized in a structure such that each operation is coordinated and can be carried out simultaneously with an another operation that cancels it out (e.g., uniting and dissociating elements, adding and subtracting; Piaget 1976a). Piaget (1976a; Piaget et al. 1992) used different mathematical and logical tools to analyze the properties of the operational structures. Piaget (1973, 1974) was very clear that the subject is only aware of the outcome (i.e., his or her performance) of these structures, but not of the structures themselves.

The operative aspect of intelligence transforms subject-object relations by inserting the data provided by the figurative functions into increasingly complex structures. In other words, the operative activity of the human mind results in the construction of more and more complex relations (spatial, causal, logical, etc.) between person and world. The operative aspect of intelligence, then, is central to understanding the kinds of qualitative changes that occur in Piaget's account of cognitive development. The figurative aspect is an auxiliary of the operative aspect of intelligence in that it provides knowledge about states that are coordinated and transformed by operations. For example, the meaning of what the subject perceives is relative to his or her action tendencies, and perceptual activities (i.e., what the subject pays attention to) are dependent on the subject's operative development (Piaget 1969).

Equilibration

At each point in development, children are in a state of equilibrium with the environment, characterized by a particular balance of assimilation and accommodation. Development is a process that leads to increasingly more stable (complete and consistent) forms of equilibrium. Piaget (1985) termed this process equilibration. Equilibration must ensure two things: (1) It must always open up new possibilities (as life is creative); (2) it must conserve previous structures as substructures in new and elaborated structures. The second requirement is necessary to account for the fact that logico-mathematical knowledge does not become invalid with the construction of new knowledge (Piaget 1972a). Equilibration thus must reconcile the two contradictory tendencies of openness and closure of structures.

The theory of equilibration takes central place in Piaget's later work (Boom 2009), in which he focused in more detail on the specific processes involved in equilibration. Although Piaget identified several processes as playing an important role in equilibration, such as dialectics, contradiction, affirmation and negation, the generation of possibilities, and the process of becoming aware, his theory of equilibration remains unfinished (Campbell 2009; Piaget 1976b, 1980, 1987). Here we focus on the role of reflecting abstraction in equilibration, as reflecting abstraction is central to the construction of more powerful knowledge structures (Piaget 1971, 1985, 2001).

Reflecting abstraction is an elaborative process by which children discover the structural aspects of their cognitive activity (Piaget 2001). For instance, putting marbles, one after the other, in a receptacle is an action with several structural aspects, one of which is based on the creation of a serial order and another on the creation of a set with a growing number of elements. By becoming aware of the relations between and coordination of their actions, children abstract structure (the coordinatory or operative aspect of actions) from content and, in turn, project this structure to a higher cognitive level.

Piaget (1950a, 1972a, 2001) suggests that the general coordinations of actions (e.g., putting things together, establishing correspondences, ordering) are the source of logico-mathematical knowledge. For example, to understand the commutativity of addition ($3+2=5=2+3$), the child needs to put down objects in a different orders (e.g., first 3, then adding 2; after that 2, adding 3) and then realize that the total remains the same (i.e., the product of the actions is independent of the particular order in which the actions are executed; Piaget and Inhelder 1976). With the emergence of the semiotic function (see below), the knowledge abstracted from the coordination of actions becomes internalized, and the commutativity of addition can be deduced by mental operations.

The mechanism of reflecting abstraction then ensures that development has an intrinsic logic and proceeds by way of successively conceptualizing the structures or forms of knowledge underlying previous knowing levels (Piaget 1971, 2001). Thus, the form of stage n becomes the content of stage n + 1. With each new and higher stage, the forms become increasingly abstract. Through the mechanism of reflecting abstraction, then, development proceeds by way of successively conceptualizing and reconstructing the knowledge structures underlying previous knowing levels, thus ensuring the generativity and rigor of knowledge.

Constructivism

Piaget's way of conceptualizing intelligence and knowledge clashes with the way it is conceptualized by empiricism. Piaget himself was ardent in his opposition to empiricism (e.g., Piaget and Inhelder 1976). According to Piaget, empiricism conceives of human beings as passive, emphasizing sense perception, which provides copies of reality, and association as major sources of knowledge. Piaget argued that empiricists misconstrue the fundamentally active relation between subject and world as a passive, causal relation. Furthermore, Piaget (1970, 1972a, b) contended that the idea that knowledge is a copy of reality is flawed because there would be no way to evaluate the accuracy of such copies because they cannot be directly compared to reality itself. Rather, he was influenced by Kant's (1929) idea that objectivity is constituted by the subject. Kant argued that our intuition (i.e., sensibility) and understanding use a priori (i.e., independent of all experience) forms and categories, which are the condition of the possibility for experiencing objectivity. Piaget (1963, pp. 376–395) subscribes to the ordering and organizing function of the mind—indeed, this is implied in the notion of assimilation. However, Piaget believed that the forms and categories are not a priori or innate but rather undergo development as a result of the subject's interaction with the world—which is implied in the notion of accommodation.

As an alternative to empiricist and nativist interpretations of knowledge, Piaget proposed a constructivist interpretation, according to which knowledge is neither a simple recording of reality nor preformed, but an active construction that "at its origin, neither arises from objects nor from the subject, but from interactions—at first inextricable—between the subject and the object" (Piaget 1970, p. 704). It is in the course of these interactions that the subject becomes aware of herself as the (social and physical) world provides resistance to her projects, and she constructs an increasingly complex knowledge of the world as she coordinates her actions and operations (Piaget 1954).

At the psychological level, constructivism amounts to a pragmatist approach to knowledge because Piaget (1970, p. 704) emphasized action as the source of knowledge: "in order to know objects, the subject must act upon them, and therefore transform them: he must displace, connect, combine, take apart and reassemble them." Action is goal-directed and, at least early in life, aims at success and not truth; it is a lived and not a contemplative intelligence (Müller 2009). Action also transforms reality itself: "The construction of an electronic machine or a sputnik not only enriches our knowledge of reality, it also enriches reality itself, which until then did not include such objects. This creative nature of action is central" (Piaget and Inhelder 1976, p. 33).

Constructivism combines genesis (empiricism) and structuralism (nativism): every genesis originates from one knowledge structure and results in another structure, and, conversely, every structure has a genesis (Piaget 1967). We now turn to the stages defined by the succession of structures.

Stages in the Development of Intelligence

In standard psychology textbooks (e.g., Berk 2012), Piaget is typically portrayed as a stage theorist who claimed that stages are general structures that define a child's behavior in each area of cognitive functioning and that age is a criterion for stage. Consequently, it is argued that Piaget's theory is flawed because empirical evidence shows that at any point in development, children's behavior is heterogeneous and not homogeneous (e.g., they may reason at a preoperational level in one conservation task, and at a concrete-operational level in another) and that particular stages emerge earlier than Piaget would predict.

This portrayal of Piaget's stage theory is utterly incorrect (Chapman 1988; Smith 1993). Piaget did not claim that stages are characterized by homogeneity, and, in fact, Piaget often made the opposite point that variability should be expected (Chapman 1988). Furthermore, variability in children's performance on structurally similar tasks is entirely consistent with the basis of Piaget's grounding assumption that thought originates in action. Based on this assumption, cognitive structures should, at first, be context and content specific. That is, cognitive structures cannot be separated from their content, and although structures involving different content (e.g., number and volume), may be of the same logical form, they develop independently in a functional sense through the child's activity with these different areas of content.

Stages are also not defined in terms of age. Rather, they are defined in terms of performance on particular tasks that Piaget analyzed in terms of the operations and structure they require. He acknowledged that the age of acquisition of operations is highly variable and influenced by the amount of cognitive stimulation. Furthermore, central to Piaget was not the age at which the stages emerge but the mechanisms involved in stage transitions. Each stage is a temporary equilibrium in the process of equilibration (Piaget 1985, p. 139). Higher stages are in a better equilibrium; they are characterized by a more complex understanding of the world (greater spatiotemporal distance) and more mobile operative structures. Because the stages build on each other, they constitute an invariant sequence. But there is no fixed, predetermined end point to the development of intelligence because a new stage always opens up new possibilities of interacting with the world, which may lead to further development. We next briefly describe the main characteristics of each stage (see Chapman 1988).

Sensorimotor Stage Piaget (1963; see Müller 2009) termed the developmental period during approximately the first 18 months of life sensorimotor intelligence. It plays a key role in bridging the gulf between the biological level of functioning and rational thought. Sensorimotor intelligence is a practical, embodied intelligence on the basis of which infants interact with the world through perception–action cycles. At the sensorimotor stage, meaning is originally embedded in unreflective activities; objects have a functional, practical meaning, they are things at hand, utensils for practical use or manipulation. Infants employ action schemes like sucking, pushing, hitting, and grasping to explore and manipulate the world. At the outset, the newborn has no self-consciousness and no clear awareness of what effects she herself produces through actions on the world and what effects occur independently of her actions. By coordinating her actions and applying them in the social domain (imitation), the infant gradually learns to distinguish between self, other persons, and world.

Piaget traced the process of differentiation and coordination of action schemes through several sensorimotor substages. Drawing on Claparède's (1917) suggestion that intelligence is an adaptation to new circumstances, Piaget highlighted the transition between sensorimotor substages III and IV. During sensorimotor substage III (approximately 4–8 months), children use secondary circular reactions (i.e., actions that focus on the effects they produce in their world) to interact with the world. Essentially, secondary circular reactions aim at reproducing the effect by repeating the action that generated this effect in the first place. For example, Piaget's daughter Lucienne moved her legs vigorously, thereby shaking her bassinet. The movement made the dolls swing that were hanging from the hood. Lucienne looked at the dolls, smiled, and repeated the movement (Piaget 1963, Obs. 94). At sensorimotor substage VI (approximately 8–12 months), by contrast, infants construct hierarchical relations between secondary circular reactions by subordinating one action as a means (e.g., removing an obstacle) to another action as an end (e.g., grasping the rattle). The coordination of secondary schemes differs from the behaviors displayed at the previous substage in two ways (Piaget 1963, p. 229). First, whereas secondary circular reactions simply tried to reproduce an interesting event, the coordination of secondary circular reactions becomes necessary when infants in pursuit of their goals encounter an obstacle that requires them to accommodate existing schemes to a new situation. Second, whereas secondary circular reactions lead to the differentiation between means and ends only after the fact, means and ends are differentiated in substage IV from the outset. At substage IV, then, children coordinate two independent schemes, the scheme assigning an end to the action (e.g., grasping the rattle) and the scheme used as a means (e.g., removing the obstacle). Because the behavior at substage IV requires an adaptation to a new situation, Piaget (1963, p. 228) terms means–end coordination a "true act of intelligence" and "the beginning of intelligent action."

The coordination and differentiation between actions result in the construction of increasingly complex relations between objects—"the objectivization of reality" (Piaget and Inhelder 1976, p. 32)—as reflected in the development of such basic categories as space, time, causality, and object (Piaget 1954). For example, in order to remove a cushion that is placed in front of an object, the child must realize for herself that the cushion, in fact, is placed in front of the object (space), that she must remove it before grasping the object (temporal series), that the object behind the cushion still exists (object permanence), and that in order to remove the cushion she must grasp it (spatialized and objectified causality).

The sensorimotor period ends with the emergence of symbol representations, which allow infants to transcend the immediate here and now. At the completion of the sensorimotor stage, for the infant, his own action is no longer the whole of reality and instead now becomes "one object among others in a space containing them all; and actions are related together through being coordinated by a subject who begins to be aware of himself as the source of actions" (Piaget 1972a, pp. 21–22).

Preoperational Stage The emergence of the symbolic or, as Piaget also termed it, semiotic function marks the onset of the preoperational stage, which extends from about 2 to about 7 years (Piaget and Inhelder 1969). The semiotic function underlies children's abilities to engage in a number of different activities, such as deferred imitation (i.e., imitation in the absence of the model), pretend play, drawing, psychological functions based on mental images (e.g., recall memory), and language. These activities are practiced and refined during the first substage of this stage, the level of preconceptual thought (approximately 2–4 years of age). At the same time, preoperational thought is characterized by profound cognitive limitations. For example, although preconceptual thought is no longer tied to particular objects or events (the here and now), it fails to distinguish between individual members of a concept and the generality of a concept. To illustrate, when Piaget's daughter Jacqueline was 31 months old, she cried upon seeing a slug, "There it is!" When she saw another slug a few yards further she said, "There's the slug again" (Piaget 1962). At this substage, concepts thus remain midway between the generality of the concept and the individuality of elements composing it. On the one hand, there is no concept of a general class; on the other hand, particular objects have less individuality and easily lose their identity.

At the second substage of preoperational thought—termed intuitive thought—symbolic representational schemes become increasingly coordinated, and children become capable of relating two such schemes to each other by means of an unidirectional logical relation (Piaget 1970). For example, in comparing the liquid in two differently shaped containers, children may use height in order to infer the amount of liquid, but ignore the width of the container. Intuitive thought thus remains centered on one dimension (e.g., height) and fails to establish bidirectional relations between dimensions (Piaget et al. 1977).

Concrete Operational Stage During the concrete-operational stage, which emerges around 6–7 years, operations (i.e., internalized actions such as putting like objects together, putting objects in one-to-one correspondence) become coordinated and integrated into logical systems (see Bibok et al. 2009). As a result, children no longer center on one aspect of a situation, and they can mentally reverse transformations that have occurred in reality. The coordination of operations into systems also leads to the emergence of logical necessity (Piaget 1976a).

Piaget devised a variety of conservation tasks to assess concrete-operational thought. Conservation refers to the understanding that a whole exists as a quantitative invariant and therefore remains intact despite the quantitative rearrangement of its parts (Piaget and Inhelder 1974). For example, the number of objects in a set does not change by rearranging them (e.g., spreading them out). To understand that the quantity has not changed, children need to coordinate transformations in two dimensions (density of objects, length of row of objects). An operative understanding of conservation is logical in nature, and it is not given by empirical observation of transformations.

Another concept that children understand at the concrete-operational level is class inclusion (Piaget and Inhelder 1969; Piaget 1980). A typical class inclusion task requires children to compare the number of objects in the including or superordinate class with the number of objects in the most numerous of two of its subclasses. For example, given 12 daisies and 4 roses, children are asked, “Are there more daisies or more flowers?” A correct answer requires that children conserve the including class (B) while making the quantitative comparison between it and the included class (A). Although this may sound simple enough, such a comparison actually involves a multistep process in which children must not only be able to construct the including class but also be able to reverse this affirmative operation by properly decomposing it. The first step involves being able to combine two subclasses to form a superordinate class, or A (daisies) + A’ (roses) = B (flowers). The second step involves performing the inverse (negative) operation associated with this combination of subclasses. This entails subtracting each subclass from the superordinate class such that A = B − A’ and A’ = B − A. The inverse operation, thus, implies that children construct each subclass through negation under the including class. Piaget termed this type of negation partial because it is applied to a part of a larger whole. Through partial negation children realize that the subclass A is an autonomous whole, which enables them to recognize that there are some B’s that are not A’s (e.g., there are some flowers that are not daisies) and that, therefore, there are more B’s than A’s. The different operations required by class inclusion provide an example of an operational structure (Piaget 1976a).

Formal Operational Stage The last stage of cognitive development described by Piaget emerges during adolescence (Inhelder and Piaget 1958; see Moshman 2009). Piaget and his collaborator Bärbel Inhelder studied formal operations by presenting children and adolescents with concrete material (e.g., different weights, strings of different length) to be manipulated in order to discover scientific laws or the cause of a result from several possible factors (e.g., which factor—weight, length of string, height of dropping point, force of push—determines the frequency of the pendulum’s oscillation). These studies revealed that children approached scientific problems in a qualitatively different way than adolescents. Although children were capable of classifying and cross-classifying the independent variables, of properly ordering magnitudes of the independent variable along one dimension, and of putting these seriations into correspondence with their effects on the dependent variable, they failed to separate the involved variables by varying only one variable and holding all others constant. As a result, these children did not supply adequate proof for their statements. By contrast, from the outset, adolescents formulated hypotheses and derived conclusions from these hypotheses. They then proceeded to test these hypotheses by systematically controlling all variables except the one under investigation. Thanks to their systematic experimental approach, adolescents excluded hypotheses that were contradicted by observations and converged on the hypothesis that was actually true.

For Piaget, the difference between children's and adolescents' approaches to these problems suggested the reversal of the direction between reality and possibility: whereas in the concrete-operational stage, possibility remains an extension of reality, in the formal-operational stage, reality is subordinated to possibility. Adolescents are capable of thinking hypothetico-deductively by drawing necessary conclusions from truths that are considered merely possible.

Semiotic Function and Intelligence

Piaget (1963; Piaget and Inhelder 1971) held that consciousness is always based on signs or, better, signifiers. Signifiers are items that convey meaning. At the sensorimotor level, signifiers are not yet differentiated from their referent (signifieds). Signifiers at this level are termed indications. An indication is an "objective aspect of external reality" (Piaget 1963, p. 193), "a perceptible fact which announces the presence of an object or the imminence of an event (the door which opens and announces a person)" (Piaget 1963, pp. 191–192). Signifieds at this level are sensorimotor schemes that confer meaning on the elements interacted with.

At the end of the sensorimotor stage, the coordination and differentiation of schemes culminate in the emergence of signifiers that are differentiated from their signifieds. Piaget (Piaget and Inhelder 1969, 1971) termed a system of such signifiers the semiotic function. The semiotic function subsumes both symbols and signs. Piaget defined symbols such as mental images as signifiers (i.e., they resemble the things signified) and signs, such as words, as arbitrary and conventional signifiers. The semiotic function makes it possible for children to form mental representations and to think about absent objects as well as past, future, and even fictitious events. It also increases the speed of processing because it makes it possible to imagine at the same time the successive phases of an action. Finally, it opens up the possibility of reflecting on and understanding the reasons why some actions are successful and others not (Piaget 1954).

During the preoperational period, children use symbols in symbolic play (e.g., a toy cup stands for a real cup), deferred imitation (e.g., imitating an action of an absent model), and drawing (Piaget 1962; Piaget and Inhelder 1971). Piaget believed that particularly young children need to rely on the use of individualized and personal systems of symbols (Piaget and Inhelder 1971). This is because personal symbols make fewer processing demands than language which is based on collective and arbitrary signs. Piaget recognized that language is essential to socialization, which, in turn, modifies action and behavior. Verbal exchange between individuals allows children to share ideas, and the resulting "collective concepts" reinforce individual thinking (Piaget 1995). Being more mobile than symbols, language also makes a unique contribution to the mobility of thought.

At the same time, neither language nor symbols are the source of the forms of thought found at the concrete- and formal-operational stages. According to Piaget (1970), these forms of thought are grounded in the practical coordination of actions (e.g., grouping objects, seriating objects) at the sensorimotor stage. The semiotic function, and particularly language, is necessary for the internalization of actions (i.e., without the semiotic function, operations would have to be executed as successive actions and could not be condensed into a simultaneous whole), but it is not sufficient to explain logical thought. In sum, Piaget considered the semiotic function only a tool used by and dependent on the operative aspect of intelligence (Piaget and Inhelder 1969, 1971).

Affectivity and Intelligence

There is a long tradition of treating intelligence and emotion as distinct. Even in present-day psychology, IQ and EQ are thought of as separate (or even opposing) constructs (Goleman 1990).

In contrast to dualistic conception, Piaget (1981; see Sokol and Hammond 2009) believed that all behaviors involve an affective aspect and a cognitive aspect. The affective aspect is responsible for motivating the organism's interaction with the environment by assigning a value or goal to the behavior. However, achieving a particular end can involve a number of different paths. It is the cognitive aspect of behavior that structures such paths and thus the relation between the individual and the environment. In other words, affect provides the values and ends for actions, whereas cognitive functions are the means for achieving the ends (see Binet and Simon 1916 p. 142).

To illustrate that any intelligent act contains both affective and cognitive contributions, take the following sensorimotor action: a child reaches for a toy by pulling on the blanket under the toy. This act has an affective component. In fact, two types of affectivity are involved in this act: synchronic affectivity (in the moment) and diachronic affectivity (over time). First, the child evaluates his current actions with feelings of success and failure (synchronic affectivity), and second, the child's evaluations of the situation involve a system of values that he has developed over time, engaging his interest in obtaining the toy (diachronic affectivity). These affective components regulate the cognitive component of the act that facilitated obtaining the object by pulling on the blanket. Thus, affect provides direction for intelligence, first by regulating interest and effort and, second, by assigning value to solutions sought. As such, for Piaget, affect and intelligence are inseparable, and Piaget (1981) often underscored the role of affectivity in intellectual growth.

Social Interaction and Intelligence

Piaget is often accused of failing to address the role of social interaction in development. This, however, is not the case (see Kitchener 2009). Piaget was by no means oblivious to the role of the social in development, as attested by, for example, his statement that "human intelligence is subject to the action of social life at all levels of development from the first to the last day of life" (Piaget 1995, p. 278). In his work, Piaget struggled with the fundamental epistemological question: "Is it the individual as such or is it the social group that constitutes the motor or, if you prefer, the 'context' of intellectual evolution?" (Piaget 1995, p. 215). Piaget contrasted his own solution to this question with theoretical positions which suggest that rationality is derived either from the individual or the collective. By reducing the social to the aggregation of ready-made individual consciousnesses, individualism provides an atomistic explanation of the social and rationality. Collectivism, on the other hand, considers the social as a whole that cannot be derived from an additive composition of individuals. Rather, the collective whole is characterized by emergent, novel properties and structures, and it modifies its members (i.e., individual persons; see Piaget 1995).

Piaget criticized both individualism and collectivism and proposed an interactive relational position as an alternative explanation of the role of social interaction in intellectual development. According to the interactive relational position, "there are neither individuals as such nor society as such. There are just interindividual relations" (Piaget 1995, p. 210). These relations between individuals are primary and "constantly modify individual consciousnesses themselves" (Piaget 1995, p. 136). The interactive relational point of view leads to a more fine-grained analysis of specific social relationships and their implications for development. Piaget describes two extreme types of social interaction: constraint and cooperation. Whereas constraint involves the imposition of authority and group traditions on the individual, cooperative interactions are based on reciprocity and equality. In his early work, Piaget (1932, 1995) argued that cooperative relations among equals are necessary for the development of rationality and autonomous morality; by contrast, relations of unilateral respect in which one individual has to submit to another individual's authority impede the development of morality and rationality.

The upshot of the interactive relational view is that because individuals must coordinate their actions vis-à-vis the world, "[i]ndividual operations and cooperations form one inseparable whole in such a way that the laws of the general coordination of actions are, in their functional nucleus, common to inter- and intraindividual actions and operations" (Piaget 1971, p. 98). Individual operations and cooperations are subject to the same kind of combinations and transformations as actions and operations, thus the question of whether rationality is essentially social or individual becomes moot:

> To wonder whether it is intrapersonal operations that engender interpersonal co-operations or vice versa is analogous to wondering what came first, the chicken or the egg … The internal operations of the individual and the interpersonal coordination of points of view constitute a single and the same reality, at once intellectual and social. (Piaget 1995, pp. 294, 307)

For that reason, social interactions are subject to and regulated by the same equilibration processes as intraindividual actions and operations.

Conclusion

In this chapter, we described, in broad strokes, Piaget's theory of intelligence. We showed that Piaget defined intelligence as the cognitive organization of an organism. More specifically, Piaget argued that intelligence can be defined functionally as resulting in increasing spatiotemporal distances between subject and world and structurally in terms of a sequence of stages that move toward increasing reversibility. We showed how Piaget's conception of intelligence is rooted in his larger epistemological framework, and we described major features of this framework.

Even though many aspects of Piaget's theory have been heavily criticized (Lourenço and Machado 1996), we believe that the way in which he conceptualized the relations between affectivity, symbols, social interaction, and intelligence was original and provides a fruitful direction for overcoming sterile dichotomies. Furthermore, we think that three features of his theory are essential to any comprehensive and coherent theory of intelligence. First, theories of intelligence must address the biological dimension of intelligence. Elegantly, Piaget's insight into self-organization grounded his conception of intelligence in the very feature of life itself. Second, intelligence is not passive, not a process triggered by an input and, in turn, triggering some sort of output. Theories of intelligence must capture the fact that human beings are active and transform through their actions the world. Intelligence is not just comprised of a set of theoretical abilities; it includes, and is grounded in, practical skills. Third, intelligence has a normative dimension. This normative dimension includes values, moral norms, and logical necessity (Smith 1993, 2009). The normative dimension, which may be unique to humans, has been recalcitrant to reductionist explanations. We submit that a successful theory of intelligence will be based on the valuable insights provided by Piaget's theory, that is, modern intelligence theories must be grounded in biology, acknowledge the active nature of intelligence and the role of practical intelligence, and capture the normative dimension of knowledge.

References

Aristotle. (1956). *De anima*. Oxford: Clarendon.

Aristotle. (1970). *Physics*. Oxford: Clarendon.

Aristotle. (1971). *Metaphysics*. Oxford: Clarendon.

Bennour, M., & Vonéche, J. (2009). The historical context of Piaget's ideas. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 45–63). New York: Cambridge University Press.

Berk, L. (2012). *Infants and children* (7th ed.). Toronto: Pearson.

Bibok, M. B., Müller, U., & Carpendale, J. I. M. (2009). Childhood. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 229–254). New York: Cambridge University Press.

Binet, A. (1894). *Introduction à la psychologie expérimentale* [Introduction to experimental psychology]. Paris: Alcan.

Binet, A. (1975). *Modern ideas about children*. Menlo Park: San Francisco State University. (Original work published 1909).

Binet, A., & Simon, T. (1916). *The intelligence of the feeble-minded*. Baltimore: Williams & Wilkins Company. (Original work published in 1909).

Boom, J. (2009). Piaget on equilibration. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 132–149). New York: Cambridge University Press.

Bringuier, J.-C. (1980). *Conversations with Jean Piaget*. Chicago: The University of Chicago Press. (Original work published in 1977).

Bühler, K. (1933). *The mental development of the child*. New York: Harcourt, Brace & Company. (Original work published 1918).

Campbell, R. L. (2009). Constructive processes: Abstraction, generalization, and dialectics. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 150–170). New York: Cambridge University Press.

Chapman, M. (1988). *Constructive evolution: Origins and development of Piaget's thought*. New York: Cambridge University Press.

Chapman, M. (1992). Equilibration and the dialects of organization. In H. Beilin & P. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 39–60). Hillsdale: Lawrence Erlbaum.

Claparéde, E. (1917). La psychologie de l'intelligence. *Scientia, 22*, 253–268.

Goleman, D. (1990). *Emotional intelligence*. New York: Bantam Books.

Hamlyn, W. (1971). Epistemology and conceptual development. In T. Mischel (Ed.), *Cognitive development and epistemology* (pp. 3–24). New York: Academic.

Hübner, J. (1999). Die Aristotelische Konzeption der Seele als Aktivität in de Anima II 1 [Aristotle's concept of soul as activity in de Anima II 1]. *Archiv für Geschichte der Philosophie, 81*, 1–32.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books. (Original work published in 1955).

Inhelder, B., & Piaget, J. (1969). *The early growth of logic in the child*. New York: Norton. (Original work published 1959).

Kant, I. (1929). *Critique of pure reason*. London: Macmillan. (Original work published in 1787).

Kitchener, R. F. (1993). Piaget's epistemic subject and science education: Epistemological vs. psychological issues. *Science and Education, 2*, 137–148.

Kitchener, R. F. (2009). On the concept(s) of the social in Piaget. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 110–131). New York: Cambridge University Press.

Lourenço, O., & Machado, A. (1996). In defense of Piaget's theory: A reply to 10 common criticisms. *Psychological Review, 103*, 143–164.

Moshman, D. (2009). Adolescence. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 255–269). New York: Cambridge University Press.

Müller, U. (2009). Infancy. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 200–228). New York: Cambridge University Press.

Natorp, P. (1910). *Die logischen Grundlagen der exakten Wissenschaften* [The logical foundation of the exact sciences]. Leipzig and Berlin: B. G. Teubner.

Piaget, J. (1932). *The moral judgment of the child*. London: Kegan Paul Trench Trubner.

Piaget, J. (1950a). *Introduction à l'èpistemologie gènètique, Vol. 1: La pensèe mathèmatique* [Introduction in genetic epistemology: Mathematical thought]. Paris: Press Universitaires de France.

Piaget, J. (1950b). *Introduction à l'èpistemologie gènètique, Vol. 2: La pensèe physique* [Introduction in genetic epistemology: Physical thought]. Paris: Press Universitaires de France.

Piaget, J. (1950c). *Introduction à l'èpistemologie gènètique, Vol. 3: La pensèe biologique, la pensèe psychologique et la pensèe sociologique* [Introduction in genetic epistemology: Biological, psychological, and sociological thought]. Paris: Press Universitaires de France.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books. (Original work published in 1937).

Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: W. W. Norton & Company. (Original work published in 1945).

Piaget, J. (1963). *The origins of intelligence*. New York: Norton. (Original work published in 1936).

Piaget, J. (1967). Genesis and structure in the psychology of intelligence. In J. Piaget (Ed.), *Six psychological studies* (pp. 143–158). New York: Random House. (Original work published in 1964).

Piaget, J. (1968). Explanation in psychology and psychophysiological parallelism. In P. Fraisse & J. Piaget (Eds.), *Experimental psychology: Its scope and method* (History and method, Vol. 1, pp. 153–191). London: Routledge & Kegan Paul. (Original work published in 1963).

Piaget, J. (1969). *The mechanisms of perception*. New York: Basic Books. (Original work published in 1961).

Piaget, J. (1970). Piaget's theory. In P. Mussen (Ed.), *Carmichael's manual of child psychology* (3rd ed., pp. 703–732). New York: Wiley.

Piaget, J. (1971). *Biology and knowledge*. Chicago: The University of Chicago Press. (Original work published in 1967).

Piaget, J. (1972a). *Psychology and epistemology: Towards a theory of knowledge*. Harmondsworth: Penguin Press. (Original work published in 1970).

Piaget, J. (1972b). *The principles of genetic epistemology*. New York: Basic Books. (Original work published in 1970).

Piaget, J. (1973). *Main trends in psychology*. London: George Allen & Unwin. (Original work published in 1970).

Piaget, J. (1976a). *The psychology of intelligence*. London: Routledge & Kegan Paul. (Original work published in 1947).

Piaget, J. (1976b). *The grasp of consciousness*. Cambridge, MA: Harvard University Press. (Original work published in 1974).

Piaget, J. (1980). *Experiments in contradiction*. Chicago, IL: University of Chicago Press. (Original work published in 1974).

Piaget, J. (1981). *Intelligence and affectivity*. Palo Alto, CA: Annual Reviews. (Original work published in 1954).

Piaget, J. (1985). *Equilibration of cognitive structures*. Chicago: University of Chicago Press. (Original work published in 1975).

Piaget, J. (1987). *Possibility and necessity, Vol. 1: The role of possibility in cognitive development*. Minneapolis: University of Minnesota Press. (Original work published in 1981).

Piaget, J. (1995). *Sociological studies*. London: Routledge. (Original work published in 1977).

Piaget, J. (2001). *Studies in reflecting abstraction*. Hove: Psychology Press. (Original work published in 1977).

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books. (Original work published in 1966).

Piaget, J., & Inhelder, B. (1971). *Mental imagery in the child*. New York: Basic Books. (Original work published 1966).

Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities*. London: Routledge & Kegan Paul. (Original work published in 1941).

Piaget, J., & Inhelder, B. (1976). The gaps in empiricism. In B. Inhelder & H. Chipman (Eds.), *Piaget and his school* (pp. 24–35). New York: Springer. (Original work published in 1969).

Piaget, J., Grize, J.-B., Szeminska, A., & Bang, V. (1977). *Epistemology and psychology of functions*. Dordrecht: Reidel. (Original work published in 1968).

Piaget, J., Henriques, G., & Ascher, E. (1992). *Morphisms and categories: Comparing and transforming*. Hillsdale: Erlbaum. (Original work published in 1990).

Smith, L. (1993). *Necessary knowledge*. Hove: Lawrence Erlbaum.

Smith, L. (2009). Piaget's developmental epistemology. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 64–93). New York: Cambridge University Press.

Sokol, B. W., & Hammond, S. I. (2009). Piaget and affectivity. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 309–323). New York: Cambridge University Press.

Taine, H. (1872). *On intelligence*. New York: Holt & Williams.

Vonéche, J., & Vidal, F. (1985). Jean Piaget and the child psychologist. *Synthese, 65*, 121–138.

Alfred Binet and the Children of Paris

11

Amber Esping and Jonathan A. Plucker

Alfred Binet was born in 1857 to a wealthy but troubled French family. His father, a physician, and his mother, an amateur artist, divorced when he was a child, and he grew up in his mother's household. Family resources afforded him an excellent private school education in Nice and later Paris. He distinguished himself in French composition, but his academic record was otherwise unremarkable (Siegler 1992; Wolf 1964). Binet's career path appeared desultory and unpromising at first. In 1878, he earned a law degree from the University of Paris, but never expressed any real interest in the field. He would later call law "a career for those without any [yet] chosen vocation" (Binet 1904a, p. 14). Next he attempted a medical degree at the Sorbonne in Paris, but he was profoundly distressed by the trauma and gore he witnessed in the operating room, to which he may have been especially sensitive owing to a childhood experience in which his father forced him to touch a cadaver.[1] He suffered an emotional collapse and dropped out of medical school at age 22 (Fancher 1985).

Following his psychological breakdown, Binet spent considerable time resting and reading among the peaceful stacks of the National Library of France in Paris. While browsing books on psychology, he discovered some ideas in which he could at last become genuinely and passionately invested. He first became intrigued by psychophysical experiments involving tests of two-point sensation thresholds, and he replicated some of the published experiments using himself and some friends as subjects. He concluded from his own results that extant theories about sensation thresholds should be modified, and he published a paper outlining his suggested corrections (Binet 1880). The article was well written and cogently argued, but unfortunately it was also fueled by naïve enthusiasm. The ideas Binet put forth as his own had in fact already been published—in much more sophisticated form—by a respected Belgian physiologist by the name of Joseph Delboeuf (1831–1896). Delboeuf responded by publishing a humiliating critique of Binet's article (Delboeuf 1880; see also Fancher 1985; Wolf 1964).

Undeterred, Binet continued to read about psychology and to publish articles independently,

A. Esping
College of Education, Texas Christian University, TCU Box 297900, Fort Worth, TX 76021, USA

J.A. Plucker (✉)
Department of Educational Psychology, University of Connecticut, Neag School of Education, 2131 Hillside Road, Unit 3007, Storrs, CT 06269-3007, USA
e-mail: jonathan.plucker@uconn.edu

[1] Binet told the story of touching the cadaver in 1911, indicating that this distressing experience had happened to one of his "friends." Compelling circumstantial evidence suggests that he was really talking about himself (see Wolf 1964, pp. 762–763).

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_11,

leading one biographer to dub him a "library psychologist" (Siegler 1992, p.180). He was most interested in English associationists like John Stuart Mill (1806–1873), Herbert Spencer (1820–1903), and Alexander Bain (1818–1903), but he was also influenced by Hippolyte Taine's (1828–1893) philosophical treatise *On Intelligence* (1870/1872) and Théodule-Armand Ribot's (1839–1916) treatments of English and German psychology (Foschi and Cicciola 2006). Mill was Binet's particular favorite, and he would later refer to him as "my only master in psychology" (Binet 1903, p. 68). Binet's second publication forwarded Mills' associationism as an all-encompassing explanation for the operations of the intellect (Binet 1883), but here again he was overconfident. Associationism was already beginning to lose its cachet, and prominent psychologists were routinely acknowledging its inability to account for motivational or unconscious influences (Fancher 1985). Binet eventually realized these deficiencies of associationism, but he never abandoned it completely. Indeed, his associationist roots would later be evident in his greatest achievement, the Binet-Simon Intelligence Scale (1905).

In 1883 Binet's independent means made it possible for him to volunteer his time assisting the eminent neurologist Jean-Martin Charcot (1825–1893) with his research at the Salpêtrière Hospital in Paris. Charcot was interested in hysteria, a baffling syndrome in which female patients experienced paralysis, loss of sensation, seizures, and memory loss, with no apparent neurological cause. Whereas some physicians attributed these cases to malingering, Charcot believed that the patients experienced these symptoms as real (Fancher 1985). Charcot's associated interest in hypnosis grew from the finding that some of the same "symptoms"—paralysis, amnesia, dramatic fits, and so on—could be induced through hypnotic suggestion. Therefore, the study of hypnosis offered promise in understanding the underlying causes of hysteria (Fancher 1985).

Binet assisted Charcot in hypnosis studies of hysterical patients. One particular responsibility was Blanche Witman, an intense and melodramatic young woman who was known in the wards as "The Queen of Hysterics." Ms. Witman could be relied upon to follow the three-stage hypnotic pattern of lethargy, catalepsy, and somnambulism that Charcot believed was definitive of major hypnotism. Charcot and Binet saw this easy susceptibility to hypnotic states and consistent pattern of responses as an indication of her underlying hysterical disorder. Indeed, Charcot believed generally that susceptibility to hypnosis was an indication of an underlying hysteria. This conclusion would later turn out to be incorrect (Fancher 1985).

In a series of related experiments with Ms. Witman, Binet and another of Charcot's assistants, Dr. Charles Féré (1852–1907), discovered that they could reverse or transfer Ms. Witman's behaviors under hypnosis simply by reversing the polarity of a large horseshoe magnet. For example, if Ms. Witman seemed to be paralyzed on her right side, they could transfer the paralysis to the left by reversing the magnet. They found that they could transform the expression of an emotion, such as sobbing, to its inverse (laughing) through the same mechanism. Binet and Féré were fully convinced by Charcot that deeply hypnotized people were not aware of their surroundings, so they did not consider the possibility that their patient might be attending to the magnet or related experimenter cues. In a series of articles (Binet and Féré 1885a, b, c), they attributed the magnet findings to the existence of complementary human emotions, akin to complementary colors which produce white or black when mixed (Fancher 1985).

Unfortunately for Binet, his first critic, Joseph Delboeuf, also had a side interest in hypnosis. Delboeuf respected Charcot and was ready to accept his theory of major hypnosis, but he was skeptical of the magnet findings—especially so when he saw Binet was one of the paper authors (Fancher 1985). Curious, Delboeuf visited the Salpêtrière hospital and saw immediately that Ms. Witman was aware of the magnet and was undoubtedly yielding—consciously or unconsciously—to the desires of Féré with whom she seemed to have a particular rapport. When Delboeuf undertook the same experiments under more carefully controlled conditions, he discovered that both the magnet findings and Charcot's theory of major hypnosis were invalid (see Delboeuf 1886). Binet was at first reluctant to accept Delboeuf's evidence, and a heated public exchange ensued (Binet and Delboeuf 1886). However, Binet eventually acknowledged the "loopholes for error" which had "pervert[ed]" interpretation of his study results (Binet 1896/1977, p. 76). This proved to be an essential lesson for Binet, and he was ever afterwards aware that psychological tests will always contain some degree of error. His willingness to acknowledge and work within the constraints of this fact has been called by one scholar "[perhaps] Binet's greatest contribution" to intelligence testing (Kaufman 2009, p. 22).

Following this major career setback, Binet left the Salpêtrière, and it took more than a year to find another position. Fancher (1985, p. 57) notes that under the circumstances, it is not surprising that "prospective employers did not come flocking to his door." However, he eventually connected with the director of the new Laboratory of Physiological Psychology at the Sorbonne, and he willingly accepted Binet's offer to work there, as he had done at the Salpêtrière, without compensation. He served there as a researcher and assistant director and ascended to become the laboratory's director in 1894. He held this unpaid position until he died in 1911 (Fancher 1985; Siegler 1992.)

Binet's early career may be characterized as a series of productive false starts. He made mistakes, to be sure, but in the process he gained skills and dispositions that would prove enormously beneficial in his future intelligence work. From the hypnosis debacle, he learned the importance of careful attention to experimental controls. But he also learned to appreciate the advantages of the detailed case study approach to research. This distinguished him from contemporaries like Francis Galton (1822–1911) and James McKeen Cattell (1860–1944), who favored generalizations based on large sample sizes. This appreciation of the uniqueness of individuals convinced Binet that measuring psychological variables was a far more complex and nuanced process than other intelligence researchers had so far been willing to acknowledge (Fancher 1985). His time at the Salpêtrière also allowed Binet's passive associationist psychology to mature into a sophisticated theory that recognized the active role of human attention, as well as the importance of innate and hereditary factors in determining one's makeup (Fancher 1985).

Binet and Experimental Child Psychology

Binet's curiosity was unbounded, and he produced many other significant acts of scholarship while working with Charcot at the Salpêtrière. His publications during this time included three books and more than 20 articles exploring a wide variety of subjects. Among these were the psychic life of microorganisms (Binet 1887a), sexual fetishes (Binet 1887b), and the nature of human consciousness (Binet 1890a). He also developed an interest in the natural sciences, eventually earning a Ph.D. for a dissertation on the anatomy and physiology of the subintestinal nervous systems of insects (Binet 1894). The birth of his daughters, Madeleine (b. 1885) and Alice (b. 1887), provided an avenue to study child psychology, and in 1890 he published three articles describing experiments he conducted using his girls and their friends as subjects (Binet 1890b, c, d). This emerging interest in the psychology of children evolved into Binet's new career as an experimental child psychologist (Fancher 1985).

Binet derived his first experiments with Madeleine and Alice from Galton and Cattell's psychophysical tests of reaction times and sensory acuity, which up to this point in history represented the state of the art in intellectual testing. He noted that on average, his young subjects reacted to stimuli much more slowly than did adults but also that they were far less consistent in their performances; a child's reaction time might be on par with the typical adult in one trial and substantially slower in the next. Binet deduced that the salient difference between adults and children, then, was not really the reaction times but rather the children's limited ability to sustain attention during the trials. This insight about the importance of attention proved to be fundamental to the eventual development of his intelligence scale (Fancher 1985).

Binet's psychophysical tests of color perception also yielded interesting results. Child subjects were much slower than adults in naming colors, and this outcome might have been used to support the hypothesis that children had less developed sensory acuity than adults. However, when Binet asked the child subjects to match colors, they were nearly as fast as the typical adult. From this Binet concluded that seeming adult-child differences in color perception were in actuality methodological artifacts resulting from differences in language development—the kids could see the differences; they just could not *say* them fast enough. Binet eventually lost faith in psychophysical testing as a reliable and valid measure of intellectual ability and determined that more complex, language-based tasks were needed to discriminate child from adult intellectual capacity (Fancher 1985).

Binet advanced his understanding of the importance of language development by asking his children and their friends to define common words. He discovered that children typically responded by providing concrete, functional examples of how the items were used rather than the abstract dictionary-type definitions most adults provided. For example, a knife was defined as "to cut meat." The definition of snail was "squash it." From this adult-child difference, Binet concluded that the ability to think in abstract terms must somehow be important to the development of human intelligence (Binet 1890b; see also Fancher 1985). He continued to study his children, retaining some of the Galtonian psychophysical tasks and also including tests of memory, judgment, imagination, and inkblot interpretation, as well as qualitative impressions about their temperaments and personalities. He published these results in book form in 1903 (Binet 1903).

Other aspects of human intellectual development also caught Binet's attention during these early years at the Sorbonne. He expanded his subject pool to include children in the local schools, undertaking studies of memory and suggestibility. He discovered that both accuracy of children's memories and their ability to resist the influence of experimenter suggestion improved with age (Binet 1900). He also initiated several in-depth case studies of people with extraordinary abilities and accomplishments, such as chess prodigies and mathematical wizards (Binet and Henneguy 1894) and eminent French authors (Binet and Passy 1895). From these, he determined some unanticipated facts about the human intellect. First, there are many ways of becoming extraordinary; the great writers and math and chess prodigies approached their cognitive tasks in a variety of ways. Second, for the most part, these extraordinary individuals were quite ordinary in domains other than one particular narrow area of excellence. Binet recognized these findings as important evidence of the complexity and heterogeneity of intellectual operations (Fancher 1985). The psychophysical testing that had dominated the field to this point would never be able to tease out these kinds of nuances. New methods for testing individual differences in intellectual functioning had to be developed.

Binet and Individual Psychology

However valuable the in-depth case analyses Binet cut his teeth on, he also recognized that these long investigations were not always practical. Psychologists needed to be able to compare intellectual functioning quickly along some standard dimensions, preferably in one sitting. His prior research had illuminated the vulnerabilities of psychophysical testing, so the relatively fast methods he sought would have to test higher-order cognitive processes. Binet and his research assistant, Victor Henri (1872–1940), identified 10 candidate variables for measurement: (1) memory, (2) imagery, (3) imagination, (4) attention, (5) comprehension, (6) suggestibility, (7) a esthetic sentiment, (8) moral sentiment, (9) muscular strength and willpower, and (10) motor ability and hand-eye coordination. The last two variables resonated with earlier psychophysical testing approaches, but as conceived they were more complex than standard tasks of that kind. The other eight variables were refreshingly original in flavor (Fancher 1985). Binet named this new approach "Individual Psychology" (Binet & Henri, 1986).

In 1899, a young medical student named Théodore Simon (1873–1961) contacted Binet and requested an opportunity to work with him. Binet did not need another assistant, and he was inclined to refuse the offer. However, Simon had recently obtained a medical internship working with approximately 300 abnormal children[2] at the Perray-Vaucluse asylum, near Paris, and Binet found the opportunity to apply Individual Psychology with this special population very attractive. He accepted the offer of help and trained Simon to use his testing techniques.

Simon returned to Perray-Vaucluse and spent the next several months engaged in psychological testing. These data would later become his doctoral thesis in medicine (Wolf 1961). Unfortunately, the results of Binet's Individual Psychology research program were largely disappointing. In a 1904 paper, Binet reported that they had failed to produce a valid and discriminating psychological test that could be administered in a short period of time. The in-depth case study, it seemed to him, was still the most promising approach to individual psychology (Binet 1904b; see also Fancher 1985). However, in short order Binet would be offered a challenge that would change his mind.

Binet Invents the Intelligence Test

By the early part of the twentieth century, French national laws had begun mandating public school education for all children, including children with mental disabilities, who had previously been excluded entirely or permitted to drop out early from schooling. In 1904, officials of the French government asked Binet to join a distinguished commission of experts who could provide insight and leadership regarding the education of these special cases. Binet's Individual Psychology research, his publication record, and his particular experience with Simon's institutionalized children uniquely qualified him for this undertaking. He immediately recognized the need for a diagnostic system that could identify those children who could benefit from special education classes and, just as important, prevent intellectually normal children from being misdiagnosed (Binet and Simon 1905a). One year later, he had one: The Binet-Simon Scale, the world's first modern intelligence test (Binet and Simon 1905b). In a 1909 book, Binet described the enthusiasm with which he approached this work:

> There is nothing like necessity to generate new ideas. We undoubtedly would have retained the status quo…if a matter of true social interest three years ago, had not made it mandatory for us to measure intelligence by the psychological method. It had been decided to try to organize some special classes for abnormal children. Before these children could be educated, they had to be selected. How could this be done?…It was under these circumstances that our devoted collaborator, Dr. Simon, and I formulated a plan for measuring intelligence. (Binet 1909/1973, pp. 104–105)

The definition of "intelligence" is a difficult thing to pin down even in the twenty-first century (Plucker and Esping 2014), and Binet and Simon were working from scratch. They began by looking for evidence of what might now be termed "face validity"—that is, by recruiting groups of children who had previously been identified by experts as being obviously intellectually normal or clearly subnormal in their intellectual functioning. Drawing on their earlier work in Individual Psychology, Binet and Simon administered a variety of tests to both groups, with the expectation that some of these tests might plainly differentiate normal from subnormal children. In choosing their tasks, the researchers were particularly careful to avoid items that might rely heavily on formal education, as they wanted their tests to show evidence of psychological functioning, not educational attainment. This remains an essential goal of intelligence testing to the

[2] The language used to describe intellectual and developmental disability in the late nineteenth and early twentieth centuries included the (now offensive) terms abnormal and feebleminded and the clinical labels moron, imbecile, and idiot. Binet preferred the term *débiles* ("weak ones"). The person-first language considered respectful in the twenty-first century (e.g., "persons with intellectual disabilities") was unheard of.

present day (Kaufman 2009). As one means of accessing higher-order processes, they chose to include some questions about typical life within the French cultural context. They believed that it was safe to assume that even poor children of normal intelligence would have reasonable familiarity with this kind of information (Fancher 1985).

Binet and Simon's first attempts at differentiating intellectually normal from subnormal children were unsuccessful. There were able to find important differences in average performance on the tasks, but they failed to find any set of items that only the normal children could solve. There was always overlap, with some normal children failing tests that some subnormal children passed. The "aha!" moment came when the researchers recognized one essential difference between the two groups: the normal children were able to respond to the tasks correctly at an earlier age than the other group. It was critical to take age into consideration when scoring (Fancher 1985).

Armed with this insight, the researchers created a series of tasks of increasing complexity. Some of the simplest test items assessed whether or not a child could follow a lighted match with his eyes, take a candy out of a wrapper, or shake hands with the examiner. Slightly harder tasks required children to point to various named body parts, repeat back a series of 3 digits, repeat from memory a 15-word sentence, and define words like *house, fork*, and *mama*. More difficult test items required children to state the difference between pairs of things, reproduce drawings from memory, and construct sentences from three given words such as *Paris*, *gutter*, and *fortune*. Some of the hardest items asked children to repeat back seven random digits, find three rhymes for the French word *obéissance*, state the difference between abstract concepts like *sad* and *bored*, and answer questions such as, "My neighbor has been receiving strange visitors. He has received in turn a doctor, a lawyer, and then a priest. What is taking place?" (Fancher 1985; Kaufman 2009).

The scale was revised in 1908 and 1911. The newer versions were developed with larger sample sizes, greater age, and socioeconomic ranges, and items calibrated such that they could be "located" at ages where typical children started to complete them successfully (Binet 1911; Binet and Simon 1908). For example, a 10-year-old child who completed all the tasks usually passed by 10-year-olds—but nothing beyond—would have a mental level that exactly matched his or her chronological age, 10.0. Children who attained a mental level 2 or more years behind their chronological age—e.g., a 10-year-old child with a mental level of 8—were generally diagnosed as being mentally subnormal, providing that they were otherwise healthy and motivated when they took the test.[3] This diagnosis was applied to approximately 7 % of the students who were tested (Fancher 1985).

The creation of the Binet-Simon Scale marked the development of a completely revolutionary approach to the measurement of human intellectual functioning. Rather than relying on simple measures of reaction time and sensory acuity, Binet and Simon's test purported to measure higher-order processes such as memory, language, and attention. In particular, however, the researchers believed that their scale measured the subjects' capacity to exercise judgment. Although conventional academic wisdom purports that Binet and Simon did not have a clear definition of intelligence guiding their work, they were rather clear about their conceptualization of the construct:

> [I]n intelligence there is a fundamental faculty, the alteration or the lack of which, is of the utmost importance for practical life. This faculty is judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. A person may be a moron or an imbecile if he is lacking in judgment; but with good judgment he can never be either. Indeed the rest of the intellectual faculties seem of little importance in comparison with judgment. (Binet and Simon 1916/1973, pp. 42–43)

[3] Binet and Simon were keenly aware that physical problems could mimic psychological ones. Their experience at the laboratory school they set up revealed that 5 % of students experienced academic problems merely because they could not see the blackboard (Binet 1907).

Mental Orthopedics

In 1905, Binet submitted a report in which he outlined his recommendations for special education pedagogy. He was optimistic about opportunities to help subnormal children improve their intelligence, and he strongly disavowed the popular notion that intelligence should be viewed as a fixed and immutable quality. He later stated the case this way:

> I have often realized, with great sorrow, the existence of frequent prejudice against the educability of intelligence. The well-known proverb that says: "When one is stupid, it is for long" seems to be taken literally by some teachers … they don't care about less-intelligent pupils, they don't nourish any liking or respect towards them … Intelligence is not a unique, indivisible function, a particular essence, but it's made up of the cooperation among all the minimal functions of discrimination, observation, retention, etc., whose plasticity and extensibility have been verified … As a consequence, intelligence is susceptible t development; through exercise, training, and, above all, method, one will be able to increase one's attention, memory, judgment, and to become literally more intelligent than before. (Binet 1909/1973, pp. 100–102)

Binet developed a series of cognitive exercises he called "mental orthopedics," which he believed could raise children's intelligence. A particular focus of these exercises was improving the subjects' capacity to pay attention, since this seemed to be fundamentally lacking in many children of low intelligence. For instance, he advocated for the use of fun games like "statue" in which children had to freeze until they were permitted to move (Binet 1909/1973; see also Fancher 1985). In 1907, he set up three experimental special education classes where mental orthopedics could be practiced. (The law mandating special education would not go into effect for 2 more years.) It is notable, however, that he also advocated for educating intellectually normal and subnormal children together. He believed that this practice would provide positive models for the slower-learning children and healthy opportunities for the faster-learning children to exercise virtues of duty and solidarity (Binet 1909/1973 Binet and Simon 1908; see also Foschi and Cicciola 2006).

The Binet-Simon Scale Comes to the United States

The ultimate popularity of the Binet-Simon Scale owes a large debt to the actions of the American psychologist Henry Herbert Goddard (1866–1957). One year after Binet and Simon published the first version of their intelligence test, Goddard accepted a position as Director of Research at the Training at a school for feeble-minded children in Vineland, New Jersey. The United States did not possess a uniform system for defining, diagnosing, and classifying intellectual disability, and most educators and physicians depended on a highly subjective and unreliable "we know it when we see it" approach. Goddard was fairly confident in his own judgment in these matters, and he was convinced that most people who worked closely with disabled persons could also be relied on to make "rather accurate" intuitive judgments (Goddard 1908b, p. 12). However, as a scientist, he would have preferred an objective method, had one been available. But the major steps recently taken in France had not yet made their way across the Atlantic (Zenderland 1998).

For the next 2 years, Goddard experimented with several unsuccessful approaches to mental testing. In 1908, he took an extended trip to Europe to seek counsel with experts there. On one of these visits, he met a Belgian physician and special educator named Ovide Decroly, who shared a copy of the Binet-Simon Scale. Intrigued, Goddard brought the test back to the United States and tried the tasks with the students at the Vineland school. He discovered that the mental levels of the children generally corresponded to the intuitive judgments made by himself and the other members of the Vineland staff, thus providing evidence of criterion validity (Goddard 1908a). Soon thereafter, the American Association for the Study of the Feeble-Minded tentatively adopted Goddard's classification system as "the most reliable method at present in use for determining the mental status of feeble-minded children" (Rogers 1910). With this adoption, Binet's approach to intelligence testing became firmly entrenched in American society (Zenderland 1998). Over the next few years, Goddard distributed 22,000 copies of his English translation of the Binet-Simon test (Fancher 1985). It is an irony of history that the Binet test did not become popular in France until the mid-1900s, when a French social worker who had spent time in the United States brought a US version of the test back to France (Kaufman 2009; Siegler 1992).

The Binet Tests and US Immigration Restriction

Between 1890 and 1910, approximately 12 million immigrants attempted to enter the United States through the Ellis Island Checkpoint. Immigration critics warned that this generation was "less educated, more impoverished, and more culturally 'alien' than earlier groups of immigrants" (Zenderland 1998, p. 263). To allay fears, Congress passed an 1882 law prohibiting "idiots" and "lunatics" from passing through the gates. The law expanded in 1907 to include "imbeciles," feebleminded persons, and persons with physical defects that might prevent them from sustaining themselves through respectable employment (Zenderland 1998).

Goddard and his team were invited to Ellis Island to help enforce these regulations; the Binet-Simon Scale proved central to his task. The procedure Goddard developed in 1912 was a two-step process: one assistant would visually screen for suspected mental defectives as the immigrants passed by (using the intuitive judgment purportedly developed through close contact over many years). Those who appeared suspect would then proceed to another location where the other assistant would test them with a variety of performance measures and a revised version of the Binet Scale. The number of immigrants who were deported increased exponentially as a result of these screening measures (Zenderland 1998).

Binet's Influence on Future Intelligence Tests

Binet contracted an illness and died in 1911 at the age of 54. His premature death cut short a prodigious career in its prime. Even so, the legacy he left is staggering in its influence. Aside from the unparalleled accomplishment of the 1905 test and its subsequent revisions, Siegler (1992) notes the importance of Binet's willingness to discuss frankly the virtues and limitations of his scale, and his progressive ideas about the malleable nature of intelligence. These remain hot topics in the present day. His careful attention to empirical evidence—learned the hard way through embarrassing experiences in Charcot's laboratory—distinguished him from contemporaries, like Goddard, who were more comfortable trusting subjective expert judgment. The Binet-Simon Scale has been translated into dozens of languages and revised and adapted countless times by intellectual heirs who appreciated the originality and utility of the tasks. Even though more recent approaches to intelligence theory and testing vary considerably in their theoretical orientations and in their approaches to testing, many of the items on Binet's original scale have stood the test of time (see, e.g., the enduring popularity of the Stanford-Binet assessments) and would seem familiar to twenty-first-century test takers and psychometricians.

References

Binet, A. (1880). De la fusion des sensations semblables. *Revue Philosphique, 10*, 284–294.
Binet, A. (1883). Le Raisonnement dans les perceptions. *Revue Philosophique, 15*, 406–432.
Binet, A. (1887a). La vie psychique des micro-organismes [The psychic life of microorganisms]. *Revue Philosophique, 24*(449–489), 582–611.
Binet, A. (1887b). Le fetichisme dans l'amour. *Revue Philosophique, 24*, 142–167.
Binet, A. (1890a). *On the double consciousness. Experimental psychological studies*. Chicago: The Open Court.
Binet, A. (1890b). Children's perception. *Revue Philosophique, 30*, 582–611.

Binet, A. (1890c). Studies on movements in some young children. *Revue Philosophique, 29*, 297–309.
Binet, A. (1890d). The perception of lengths and numbers in some small children. *Revue Philosophique, 30*, 68–81.
Binet, A. (1894). *Contribution à l'étude du système nerveux sous-intestinal ds insects.* [A contribution to the study of the subintestinal nervous system of insects.]. Paris: Alcan.
Binet, A. (1896/1977). Alterations of personality. In D. W. Robinson (Ed.), *Significant contributions to the history of psychology* (Vol. V). Washington, DC: University Publications of America.
Binet, A. (1900). *La suggestibilité*. Paris: Schleicher.
Binet, A. (1903). *L'Etude experimentale de l'intelligence*. Paris: Schleicher.
Binet, A. (1904a). La création littéraire: Portrait psychologique de M. Paul Hervieu. *L'Année Psychologique, 10*, 1–62.
Binet, A. (1904b). Sommaire des travaux en cours a la societe de psychologie de l'enfant [Summary of the work in progress at the society of child psychology]. *L'Année Psychologique, 10*, 116–130.
Binet, A. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'Année Psychologique, 12*, 191–244.
Binet, A. (1907). La valeur medicale de l'examen e la vision par les instituteurs. *Bulletin de la Societe libre pour l'etude psychologique de l'enfant, 40*, 146–163.
Binet, A. (1909/1973). *Modern ideas about children* (trans: Heisler S.). Menlo Park, CA: Suzanne Heisler.
Binet, A. (1911). Nouvelles recherches sur la mésure du niveau intellectuel chez des enfants d'école. *L'Année Psychologique, 17*, 145–201.
Binet, A., & Delboeuf, J. L. R. (1886). Les diverses écoles hypnotiques. *Revue Philosophique, 22*, 532–538.
Binet, A., & Féré, C. (1885a). L'hypnotisme chez les hystériques: Le transfert. *Revue Philosophique, 19*, 1–25.
Binet, A., & Féré, C. (1885b). Hypotisme et responsabilité. *Revue Philosophique, 19*, 265–279.
Binet, A., & Féré, C. (1885c). La polarization physique. *Revue Philosophique, 19*, 369–402.
Binet, A., & Henneguy, L. (1894). *La psychologie des grands calculateurs et jouers d'écecs*. Paris: Hachette.
Binet, A., & Henri, V. (1896). La psychologie individuelle. *L'Année Psychologique, 2*, 411–465.
Binet, A., & Passy, J. (1895). Nots psychologiques sur les auteurs dramatiques. *L'Année Psychologique, 1*, 60–118.
Binet, A., & Simon, T. (1905a). Sur la necessité d'établir un diagnostic scientifique des états inférieurs de l'intelligence. *L'Année Psychologique, 11*, 161–190.
Binet, A., & Simon, T. (1905b). Méthodes nouvelles pour le diagnostic du niveau intellectual des anormaux. *L'Année Psychologique, 11*, 191–244.
Binet, A., & Simon, T. (1908). The development of intelligence in the child. *L'Année Psychologique, 14*, 1–90.
Binet. A., & Simon, T. (1916/1973). *The development of intelligence in childre*n. Baltimore/New York: Williams & Wilkins/Arno Press.
Delboeuf, J. L. R. (1880). Note in *Revue Philosophique, 10,* 644–648
Delboeuf, J. L. R. (1886). Influence de l'education et de l'imitation dans le somnambulisme provoqué. *Revue Philosophique, 22*, 146–171.
Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York: W. W. Norton & Company.
Foschi, R., & Cicciola, E. (2006). Politics and naturalism in the 20th century psychology of Alfred Binet. *History of Psychology, 9*(4), 267–289.
Goddard, H. H. (1908a). The Binet and Simon tests of intellectual capacity. *Training School Bulletin, 5*, 3–9.
Goddard, H. H. (1908b). The grading of backward children. *Training School Bulletin, 5*, 12–14.
Kaufman, A. S. (2009). *IQ testing 101*. New York: Springer.
Plucker, J. A., & Esping, A. (2014). *Intelligence 101*. New York: Springer.
Rogers, A. C. (1910). Editorial: The new classification (tentative) of the Feeble-Minded. *Journal of Psycho-Asthenics, 15*, 70.
Siegler, R. S. (1992). The other Alfred Binet. *Developmental Psychology, 28*(2), 179–190.
Taine, H. (1870/1872). *On intelligence.* New York: Holt & Williams.
Wolf, T. H. (1961). An individual who made a difference. *American Psychologist, 16*(5), 245–248.
Wolf, T. H. (1964). Alfred Binet: A time of crisis. *American Psychologist, 19*(9), 762–771.
Zenderland, L. (1998). *Measuring minds: Henry Herbert Goddard and the origins of American intelligence testing*. Cambridge: Cambridge University Press.

From Psychometric Testing to Clinical Assessment: Personalities, Ideas, and Events That Shaped David Wechsler's Views of Intelligence and Its Assessment

12

Mark Benisz, Ron Dumont, and John O. Willis

To many students of psychology, the name David Wechsler is synonymous with cognitive testing. Although Dr. Wechsler died over 30 years ago, new editions and revisions of the tests derived from his original Wechsler-Bellevue Intelligence Scale (W-B; Wechsler 1939) continue to list him as the sole author (e.g., Wechsler 2008, 2012, 2014) and new, but related, tests list him as senior author (e.g., Wechsler et al. 2004; Wechsler and Naglieri 2006). Wechsler's contributions to the field of psychometrics and intelligence testing go well beyond the assessment tool he first designed over 70 years ago. His test products are not only among the most widely used and most widely studied measures of cognitive ability (Wahlstrom et al. 2012), but they have also been translated into over a dozen languages and standardized on many different societies and cultural groups (Georgas et al. 2003). Despite the fact that many different reliable and valid cognitive assessment tools are available for sale, and even though significant advances have been achieved in the field of neuropsychology, the Wechsler products continue to retain their popularity (Kaufman 2009; Boake 2002).

Why do psychologists continue to prefer Wechsler scales over other tests? Certainly, the tests have excellent and evolving psychometric qualities, a wealth of research and commentary, and a valuable history of interpretation. Ironically, though, one reason for the resistance to changing from such heavy reliance on the Wechsler scales is likely because of the changes Wechsler initiated in the way contemporary psychologists think about and use tests. He was a strong advocate of elevating the tests from a purely psychometric approach to a more clinical approach (Kaufman 2009). To understand how Wechsler achieved this change, it is necessary to understand Wechsler from a historical perspective. Like other great psychologists, Wechsler was influenced by researchers who preceded him and by the dominant schools of thought that typified and influenced the culture of his time. As is often the case in science, serendipitous events also played a role in shaping Wechsler's views of psychological assessment. Wechsler has, in turn, had a profound impact on the manner in which modern

M. Benisz, Psy.D. (✉)
Kiryas Joel School District,
1 Dinev Road, Monroe, NY 10950, USA
e-mail: mbenisz@kjsd.k12.ny.us

R. Dumont, Ed.D.
Department of Psychology, Fairleigh Dickinson University (TWH101), 1000 River Road,
Teaneck, NJ 07666, USA
e-mail: dumont@fdu.edu

J.O. Willis, Ed.D.
Department of Education, Rivier University,
419 Sand Hill Road, Peterborough, NH 03458, USA
e-mail: jwillis@rivier.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_12,

psychologists assess intelligence. Many of his views on intelligence were forward thinking and continue to be relevant today.

People Who Influenced Wechsler

In the United States, at the turn of the twentieth century, the field of psychology began to change from a science that studied consciousness to a more practical science that mirrored the ambitions and attitudes of a pragmatic American society. Much attention was focused upon the new functionalist school of psychology that shifted the considerations of psychologists away from structuralism – the study of the structure of the mind, or what the mind is – to functionalism – the study of the functions of the mind, or what the mind does (Schultz and Schultz 2012). The functionalist school was directly related to the birth of applied psychology (which itself would give rise to clinical psychology), which took psychology out of the laboratory and placed it squarely into the real world. To examine the functions of the mind, psychologists focused on the study of memory, perception, feeling, imagination, and judgment. In order to successfully study these mental activities, psychologists of the time needed to develop and then use objective assessment methods.

One major influence on Wechsler's professional as well as personal life was James McKeen Cattell. Cattell was the first professor of psychology in the United States and, through his work, helped establish psychology, once regarded as a minor field of study, or in some cases a pseudoscience, as a legitimate science.

Cattell was interested in studying human abilities and combined his early laboratory training in Germany with his interest in statistical analysis, to gather large amounts of data he could tabulate and subject to statistical scrutiny (Cattell 1890). At the turn of the nineteenth century, Cattell published the results of his study conducted with Livingston Farrand in which they assessed 100 students enrolled at Columbia University on an array of measures that included tasks of sensorimotor effects, reaction times, and pressure thresholds (Cattell and Farrand 1896). Although at the time Cattell used the term *mental tests* to describe his measurements, these tests were, for the most part, quite different from tests that would later become the core of modern-day cognitive assessment. However, at least two of the tests, a test of memory, where the students had to remember and repeat strings of digits, and a timed "test of automaticity" that required the students to name colors, are very similar to tasks found on modern cognitive tests. Along with developing the tests themselves, Cattell and Farrand also gathered data on each of the students' cultural and socioeconomic backgrounds, as well as aspects related to their physical health. Cattell's early work on test development and standardization gave rise to the field of intelligence testing and must have made quite an impact on Wechsler who studied under Cattell at Columbia University.

In 1917, Cattell was fired from Columbia University for opposing the United States' conscription policy during World War I. Years later, Cattell successfully sued Columbia University, and in 1921, he used the money from that settlement to start The Psychological Corporation, a company dedicated to finding solutions to industrial problems via the science of applied psychology (Pillsbury 1947). One can only speculate on the influence that Cattell's difficult decision with regard to involvement in World War I might have had on Wechsler's own decision to apply for conscientious objector status when his time came to serve in the Army. Later, after graduating from Columbia, Wechsler would turn to Cattell to seek employment at The Psychological Corporation, the same company that would eventually publish Wechsler's intelligence scales.

Another major influence on Wechsler's professional life was Robert Sessions Woodworth, a student of Cattell's and a scholar of psychology who also taught at Columbia University. Like Cattell, Woodworth researched mental tests and in 1911, together with Frederic Lyman Wells, developed a version of the Substitution Test (Woodworth and Wells 1911), a form of which is familiar to the present-day psychologist as the

Digit Symbol, now Coding, subtest on the Wechsler scales (Wechsler 1944). In 1916–1917, Woodworth delivered a series of lectures which were later published under the title *Dynamic Psychology* (Woodworth 1918). Woodworth explained that his system of dynamic psychology was concerned with the cause and effect relationships of human behavior. He wanted to understand how humans learn and "what leads them to think and act as they do" (Woodworth 1918, p. 34). In the summer of 1917, Woodworth served as chief psychological examiner at the naval base in Brooklyn, NY, where his unit was responsible for administering group and individual tests to over 1,000 men as part of a trial that would eventually lead to the wide-scale testing of soldier recruits during World War I (Yerkes 1921). Of all the people who shaped Wechsler's persona and destiny, Woodworth is perhaps the most important for three reasons. First, Woodworth was Wechsler's mentor at Columbia University, and Wechsler completed both his master's degree and his doctorate under Woodworth's tutelage. Second, it is likely that Woodworth helped secure Wechsler a position in the Army as a psychological examiner during World War I (Wasserman 2012). Third, it is likely that Woodworth steered Wechsler away from any aspirations of employment in academia.

Cattell's desire to design a meaningful and purposeful test was achieved a little more than a decade later by Alfred Binet. In 1904, in the Republic of France, the Minister of Public Instruction sought to find a means by which children who were intellectually impaired could be excluded from public education and moved to special classrooms (Binet and Simon 1916b, p. 9). In order to assist in this endeavor, Binet and his colleague Théodore Simon developed and published in 1905 and revised in 1908 a set of tasks, arranged in order of difficulty, that were administered individually to children. The test produced a score called the "mental age" (MA) that was compared to the child's actual chronological age (CA). If the derived mental age was the same as the child's chronological age, the child was considered to be average. If MA was higher than CA, the child was considered to be advanced. If the MA was lower than the child's CA, the child was considered to be behind or "retarded." Many of the items on the original and revised Binet-Simon test are familiar to psychologists today. These include tasks such as unfinished pictures, where the child has to find the important part that had been omitted in a picture, and repetition of figures, in which the child repeats a series of dictated digits, the length of which varies according to the age of the child. The Binet-Simon Intelligence Scales also included a test which required the child to define words, and another that required the child to make change from 20 sous, a type of arithmetical reasoning. Each of these tasks is similar in nature to what Wechsler later included in his scale as the Picture Completion, Digit Span, Vocabulary, and Arithmetic subtest (Boake 2002). Binet and Simon's concept of intelligence was how well a person could adapt to a particular circumstance (Binet and Simon 1916a). This idea is reflected in the types of tasks they chose for their test.

The Binet-Simon test was brought to the United States in 1908 by Henry Goddard (1916), who arranged for the test to be translated into English. An American revision of the Binet-Simon scale was developed in 1915 by Robert M. Yerkes et al., and was called the Yerkes-Bridges Point Scale. Instead of grouping the tests by age level as in the Binet-Simon, the test items were consolidated into subtests and arranged in order of difficulty, with the examinee able to earn points for each correct response. The point scale model was later adopted by Wechsler for use in his tests (Kaufman 2009). In his role as Chief of the Section of Psychology in the United States Army during World War I, Yerkes, had an indirect impact upon Wechsler's future career as a developer of psychological tests.

In 1916 Lewis Terman, a professor at Stanford University and a pioneer in educational psychology, published the *Stanford Revision of the Binet-Simon Scale*, now universally known as the "Stanford-Binet" test. He adopted William Stern's (1912) suggestion that mental age divided by chronological age be used for an intelligence quotient or IQ. The Stanford-Binet was an extremely popular test for many decades (Louttit and Browne 1947) and in its current edition is still in use today. Terman believed that a good intelligence test could accurately assess abstract thinking, and therefore he emphasized the use of verbal and language-based tests of arithmetical reasoning and abstract thinking (Terman 1921).

Two other personalities that helped shape Wechsler's view on intelligence were Charles Spearman and Louis Thurstone. Spearman is credited with having constructed the first theory of intelligence (Kaufman 2009). Spearman (1904) theorized that individual intelligence could be represented by a unitary factor that he called *g*. He was able to demonstrate that different measures of mental abilities were all positively correlated, which indicated that there was an underlying global factor that represented intelligence. Spearman also theorized that certain mental tests would be better at measuring *g* than others. Wechsler referred to Spearman's discovery of *g* as "one of the great discoveries of psychology" (Wechsler 1944, p. 6). Wechsler's viewpoint as to the importance of *g* did not change over time. Kaufman (2009, p. 45) writes that when Wechsler visited him in 1975, he told Kaufman's students that "nothing is more important than *g* for understanding intelligence. Global ability is *the* ability that underlies my IQ tests."

In contrast to Spearman's unitary intelligence factor theory, Thurstone, a pioneer in the fields of psychometrics, believed that intelligence comprised seven independent groups factors (Guilford 1972). Thurstone's work in factor analysis led him to formulate a model of intelligence centered around "Primary Mental Abilities" (PMAs) (e.g., Thurstone 1938). These included verbal comprehension, word fluency, number facility, spatial visualization, associative memory, perceptual speed, and reasoning. Thurstone's model influenced hierarchical models of intelligence and would later influence the way Wechsler defined intelligence (Matarazzo 1972)

Wechsler's Early Life

Wechsler's early life did not presage his career, and he would have seemed an unlikely candidate for fame. The youngest of seven children of Jewish parents, he was born in the Kingdom of Romania in 1896, during a time of severe economic depression, which led to the enactment of anti-Semitic laws and the persecution of Jews (Matikainen 2006). At the time, Jewish residents of Romania were not eligible for Romanian citizenship and were thus excluded from many vocational occupations which were open only to citizens (Iordachi 2002). Even Jewish children were not immune from discrimination. As noncitizens they were barred from attending public schools unless their parents paid prohibitively expensive tuition fees. These conditions led to a mass exodus of Romanian Jews, who left the land of their birth for more tolerant countries. Between 1900 and 1902, Wechsler's parents and siblings made their way to the United States and found refuge in New York City. Before Wechsler's 11th birthday, both of his parents had died, and he was raised by his brother, Israel Spaner Wechsler, who would himself later become a renowned neurosurgeon. Wechsler attended City College of New York, earning an A.B. in 1916, and did his graduate work at Columbia University, earning a master's degree in 1917 and a Ph.D. in 1925. His first published article (his master's thesis) was a study of memory in patients with Korsakoff's syndrome (Wechsler 1917a). In his article, Wechsler lists the tests he used to evaluate the participants in his study. What is noteworthy is that Wechsler used a test he refers to as "auditory memory span for digits" and the Knox Cube Test (a nonverbal memory test which employs sequential taps on four blocks), which he refers to as a "visual memory span for movement" (Wechsler 1917a, p. 416). Versions of the auditory Digit Span are still found on the Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V; Wechsler 2014) and the Wechsler Adult Intelligence Scale, Fourth Edition (WAIS-IV; Wechsler 2008). A cube-tapping visual memory test (long absent from Wechsler scales) is included in the WISC-IV integrated (Wechsler et al. 2004) and in the Wechsler Nonverbal Scale of Ability (WNV; Wechsler and Naglieri 2006). It is also interesting to note that in the Korsakoff syndrome study, Wechsler performed a "qualitative" (Wechsler 1917a, p. 417) as well as a quantitative "analysis of errors" (Wechsler 1917a, p. 423). This utilization and integration of separate measures of verbal and nonverbal abilities, coupled with an interpretive approach that combined both qualitative and quantitative analyses, would feature prominently in Wechsler's assessment tools several decades later.

World War I and Psychological Testing in the Army

In 1917 at the beginning of World War I, a meeting of the Society of Experimental Psychologists was held at Harvard University. Yerkes, then president of the American Psychological Association, addressed the attendees. His goal was to see how American psychologists could unite and contribute to the war effort (Yerkes 1921). One of the identified needs of the Army was the assessment of new recruits in order to know which tasks they were capable of performing. Individual testing was not a practical solution for the Army as it required significant amounts of time as well as highly trained test administrators. A protocol for testing large groups of recruits needed to be developed. After receiving the necessary approvals from the Army, Yerkes assembled a team of experts to develop a battery of tests that could be administered to large groups of men. Most of the tests were adapted from existing assessments. The team evaluated tests using criteria such as speed of testing, speed of scoring, minimal amount of writing, and (perhaps surprisingly for the time) the test's indifference to the amount of schooling a recruit had previously received. Yerkes (1921) credited Arthur S. Otis for devising and adapting the majority of these tests. The tests were tried out several times, and group scores were correlated with individual scores. The correlation between group and individual tests was reported to be 0.5, which led Yerkes to decide that group examination was acceptable. From these trials, two group administered tests were developed: the Army Alpha, designed for English-speaking recruits, and the Army Beta, administered to recruits who were illiterate or had limited proficiency in the English language. If a soldier was unable to pass the group examination, he was administered an individual test such as the Stanford-Binet, the Yerkes-Bridges Point Scale, or a test that had been developed by Yerkes' team called the Army Performance Scale. This testing project, which would eventually result in the assessment of over 1.7 million soldiers, including over 83,500 individual assessments required a large number of men with knowledge of psychological testing to administer these tests (Yerkes 1921).

On June 5, 1917, the day before Wechsler's graduation with a master's degree from Columbia University, the US government required all men between the ages of 21 and 31 to register for the draft to fight in World War I. Wechsler, who had turned 21 in January of that year, was no exception. Wechsler's war-time service is not only noteworthy for how it influenced his later career, but it is also somewhat controversial. Wasserman (2012) suggested that Wechsler tried to evade the war-time draft and notes that "Given his efforts to avoid military service in 1917, it might be considered ironic that the skills he acquired and contacts he made during his military service would shape his career in assessment and test development" (p. 33). Specifically regarding Wechsler's draft status, Wasserman wrote:

> After the U.S. Congress declared war on Germany in April, 1917, Wechsler (1917b) completed his required registration for the draft, listing himself as a "Consciencious [*sic*] Objector" and as an alien who was a citizen of Romania, who was disabled by "Near Sightness" [*sic*] and "Physical Unfitness." To the item asking about his occupation he wrote "Am student in school of philosophy." Wechsler's draft registration thus used multiple methods to avoid being drafted—claiming status as a conscientious objector, claiming exemption from military service by reason of alien (noncitizen) status, and claiming physical deficiencies that would disqualify him for military service. . . . Wechsler's status

> as a noncitizen native of Romania, some 15 years after his arrival in the United States also put him at risk. As an alien, he could not be drafted, but he could be deported. (2012, pp. 31–32)

Wasserman based these statements on information derived from a copy of Wechsler's draft registration card (Wechsler 1917b). Careful examination of the draft card and of the selective service laws in effect at that time does not fully support Wasserman's assumptions. Although he was not a citizen of the United States, Wechsler was most certainly subject to the draft under the existing draft laws. The Selective Service Regulations of 1917 (U.S. Office of the Provost Marshal General 1918, p. 52) stated that "when an alien has declared his intention to become a citizen, regardless of how long ago, he is still liable to draft, even though he has not in the meantime applied for final papers." In 1917, the process for becoming a US citizen had two steps. First the alien would declare his intention to become a citizen. This process was known colloquially as filing "first papers" (DeSipio 1987). Several years later, the alien would file a petition for naturalization, so-called final papers. On this point, it is interesting to note that on the Wechsler-Bellevue Intelligence Scale Form II (Wechsler 1946), Question 8 of the Comprehension subtest is, "Why does the United States require that a person wait at least two years from the time he makes application until the time he receives his final citizen papers?" Listed as one of the highest score answers is, "Gives them a chance to prove their sincerity and desirability."

It is almost certain that Wechsler had already filed his declaration of intent before the war broke out. Item 4 on Wechsler's draft registration card asks: "Are you (1) a natural-born citizen, (2) a naturalized citizen, (3) an alien (4) or have you declared your intention (specify which)?" The answer that is written on the card is "1st paper." Further proof that he had declared his intention to become a citizen can be found in his petition for naturalization, filed one day after Wechsler entered the Army on May 12, 1918 (Wechsler 1918). On the form, it notes that Wechsler had previously filed a petition for naturalization with the "Supreme Court of Kings County, Brooklyn, NY" on March 6, 1918. Since this civilian petition or "final papers" could only be filed three years after filing first papers, it seems likely that Wechsler filed his declaration of intent sometime around 1915 and thus became subject to US draft regulations in 1917. Wasserman also made the assumption that Wechsler filled out his own registration card. This does not seem likely. The entire draft card appears to have been completed by the registrar. The draft card comprised two sides, a front and a back. The back part of the form is titled "Registrar's Report," filled out by the registrar; it included observations that the registrar made about the prospective recruit. It is here that the registrar, not Wechsler, noted that Wechsler had possible disqualifying disabilities such as nearsightedness and physical unfitness. It seems likely that these were intended to be the objective observations of the registrar, as the registrar had to certify that his own answers were true.

One of the answers on the card can be misread easily because of the handwriting, which may be why Wasserman assumed that Wechsler filled out his own form. For item 7 "What is your present trade, occupation or office?" Wasserman noted that Wechsler answered "am student in school of philosophy" (Wasserman 2012 p. 31). Upon careful examination, we believe that the answer actually reads "A.M. Student in school of Philosophy" where A.M. is likely an abbreviation of *artium magister*, or master of arts—the degree that Wechsler was completing at the time of the registration. Wechsler, in fact, was awarded the degree the very next day (June 6, 1917) at the Columbia University commencement ceremony. Furthermore, a careful examination of Wechsler's signature on the card shows that it does not match the cursive handwriting used to complete the rest of the document, and, additionally, the color of the ink used for the signature seems different from the rest of the form. It is possible that Wechsler was a bona fide conscientious objector, willing to serve his adopted country but requesting to do so in a noncombatant role. It is equally possible that Wechsler claimed to be a conscientious objector in order to escape the possibility of live combat. There is, however, no evidence that favors either of these possibilities nor is there any evidence that Wechsler tried to avoid (noncombatant) military duty.

Wechsler did enter the Army in 1918 and while waiting induction at Camp Yaphank in Long Island, NY received some preliminary training in the administration and scoring of the Army Alpha Test (Matarazzo 1972). In May of 1918, he attended basic training at Fort Greenleaf, Georgia, and in August was assigned to Fort Logan, Texas, where he was part of the psychology unit (Yerkes 1921). His responsibilities at Fort Logan consisted primarily of administering individual psychological tests, one part of the mass psychological examining of recruits that occurred during World War I. It is in the Army that Wechsler's main ideas about the nature of psychological evaluation were formulated (Boake 2002).

Postwar Experiences and Training

At the end of the war, while soldiers stationed in Europe waited to return home, some were able to participate in higher education courses such as law and medicine. This program was known as the American Expeditionary Forces (AEF) University (Cornbise 1997). Although it is unclear exactly how, Wechsler was able to take advantage of this program. He transfered to France and later to the United Kingdom, where he worked with notable contributors to the emerging science of human intelligence such as Charles Spearman and Karl Pearson, both of whom are renowned for their research into statistical analysis.

Wechsler obtained a 2-year fellowship to the Sorbonne in France and there studied the "psychophysiology of emotions" (Boake 2002). He returned to the United States in 1922 and completed his doctorate at Columbia University in 1925 with Woodworth as his mentor. Much of the research that Wechsler had done at the Sorbonne was used in his doctoral dissertation. One of the research issues Wechsler sought to explore was whether there was evidence for the existence of a general emotional factor, similar to the general intelligence factor identified earlier by Spearman. Wechsler, however, did not find any evidence to support this factor (Matarazzo 1972).

After graduating from Columbia, Wechsler was unable to find steady employment. We can speculate that some avenues of employment might not have been open to Wechsler because of his religion. In those years, careers in academia were not always open to Jews (Schultz and Schultz 2012). Even Wechsler's mentor, Robert S. Woodworth, was not immune to the prejudices of his times. In 1929, Woodworth made it clear to a Jewish student named Daniel Harris that Harris could not become his assistant on account of Harris's religion. Woodworth advised Harris to seek career opportunities outside of academia (Winston 1996).

It is interesting to speculate about how different the history of intelligence testing might have been had Wechsler been able to obtain a career in academia. Perhaps Wechsler might have settled into a long, distinguished professorship with many publications but nary a test to his name.

After short stints at various locations (including 5 years in private practice), Wechsler eventually found employment at The Psychological Corporation as an unsalaried employee. During his time there, he performed a study, funded by the newspaper the *New York World*, that measured the intelligence of women, using a sample of chorus girls (Edwards 1974). He also developed a test that measured the intelligence and alertness of taxi drivers for the Yellow Cab Company of Pittsburgh. The test employed mechanical instruments that Wechsler developed himself. Wechsler had already patented a machine he called the *photogalvanograph* that measured the "variations in the electrical conductivity of the human or animal skin every time the individual is subjected to an emotion" (Wechsler 1924). Wechsler criticized those who would use the photogalvanograph for determining a person's guilt of a crime (Jones and Wechsler 1928). At the time, some utilized the machine with a test that included reading a list of words to a suspect. The list was purposefully designed to include some words that were meant to trigger an emotional response. Jones and Wechsler were able to show that the test results lacked reliability if the procedures were not standardized properly, including the careful location of the trigger words in the list. Wechsler's forward-thinking approach to test validity and reliability would come to good use when he developed his own intelligence test.

The Development of the Wechsler-Bellevue Intelligence Scale

Wechsler was hired in 1932 by Bellevue Hospital in New York, where he eventually became the hospital's chief psychologist, remaining until he retired in 1967. In 1932, Wechsler wrote a short paper describing what he saw as the advantages of the Army Alpha test over the Stanford-Binet, namely, the examiner's ability to analyze the subtests of the Army Alpha test to determine individual "special abilities and disabilities". Here is the first time that analysis of strengths and weaknesses is mentioned. In 1935, Wechsler published a book he considered to be one of his best works, *The Range of Human Capacities*. In this book, Wechsler used Army Alpha data among other sources to argue that abilities peaked at a certain age and then began to decline. He also argued very strongly that psychologists had overestimated the range of variations among individuals and that human beings were actually surprisingly similar.

By the time Wechsler began developing his first intelligence test during the 1930s, he had identified several key problems with existing tests that he felt needed to be addressed. He believed that the existing tests were heavily loaded on verbal items that sometimes produced scores that did not reflect the real-life intelligent functioning of the examinee. Wechsler was aware of some of the problems that existed with earlier tests because of his first-hand experience testing soldiers during World War I. In 1935, Wechsler argued that the mental age or IQ assigned to an examinee often did not accurately describe the functionality of that in individual in real life. Wechsler understood that a person could score poorly on a test and yet function adequately in society. Wechsler related an early experience he had in the Army in which he evaluated a 28-year-old, white, Oklahoman soldier who had failed the Army Alpha and Beta tests. He administered both a Stanford-Binet and a Yerkes-Bridges Point Scale, and the man obtained a mental age of eight yet was able to function perfectly well as a soldier in the Army. Before joining the Army, the man worked as an oil driller, earning enough money to support his family. Wechsler also stated that the tests were completely inadequate in appropriately measuring the true mental abilities of foreign-born adults and African-Americans (1935a).

The Wechsler-Bellevue Intelligence Scale was published in 1939 and it contained several innovative features (Kaufman 2009). Wechsler used standard scores, which he referred to as deviation IQ scores, acknowledging that he borrowed this concept from other tests (Wechsler 1949a). He set the mean score, somewhat arbitrarily (Kaufman 2009) at 100 with a probable error of 10 or standard deviation of 15, which was (intentionally or not) fairly close to the middle of the range of standard deviations of the ratio IQs found on the revised Stanford-Binet (Terman and Merrill 1937; see McNemar 1942). The deviation IQ score offered several distinct advantages over the existing ratio IQ scores. As opposed to the ratio IQ score, the statistical meaning of a score on Wechsler's test did not vary from year to year. For example, a score of 115 would be in the 84th percentile at any age and a score of 85 would always be in the 16th percentile. If an examinee maintained roughly the same ability level on the test compared to peers, the examinee's score would not vary just as a function of age. The score was meaningful, regardless of age. Deviation IQ scores were also better suited to measure developing cognition in children as cognition does not increase uniformly as children age (e.g., the difference between the average cognition of a 1-year-old and 3-year-old is greater than the difference between a 15-year-old and an 18-year old). Another advantage of the deviation IQ was that mental ages and ratio IQs were especially meaningless for adults and required selecting an arbitrary maximum chronological age to use for older examinees and ways of designating mental ages above that level.

Wechsler also made use of the subtest point-scale method that had appeared on the Yerkes-Bridges Point Scale as opposed to the age scale that appeared on the Stanford-Binet. The Wechsler-Bellevue also corrected the overemphasis on verbal tests that appeared on tests the Stanford-Binet and similar tests. Wechsler weighted the verbal and nonverbal (performance) tests more or less equally (Kaufman 2009). Finally, Wechsler used census data to create an unbiased normative sample matching age, gender, and occupation to US census data (Wechsler 1944). Although his urban sample could not be matched exactly with rural occupations, Wechsler used Yerkes's (1921) report to find the IQs of people with specific urban occupations that were similar to the IQs of agricultural workers. This allowed Wechsler to substitute urban occupations for rural ones. For the time, these methods of standardization and norming, carried out personally by Wechsler and colleagues without the support of a publisher, were quite advanced (Kaufman 2009).

The Wechsler-Bellevue was divided into two scales, a Verbal scale and a Performance scale. The subtests on the Verbal scale were Information, Comprehension, Arithmetic, Similarities, and Digit Span, with Vocabulary serving as an alternate test. The subtests on the Performance scale were Digit Symbol, Picture Completion, Block Design, Picture Arrangement, and Object Assembly.

Almost all of the tasks that Wechsler chose for his original test were in fact adapted or adopted from other tests; primarily those Wechsler was intimately familiar with from his experiences as an examiner in the Army. Kaufman (2009) notes that "the similarity of Wechsler's original set of subtests to the tasks used to evaluate recruits, soldiers, and officers during World War I is striking." In fact, Wechsler (1944) stated this explicitly. He did not even change the names of many of the tests he borrowed. Wechsler's goal in developing the Wechsler-Bellevue was not to create new subtests or tasks but rather to select the best available measures of intelligence and combine them into a standardized, norm-referenced battery. Wechsler (1944) gives the source or sources for most of his tests. For example, the Information and Comprehension subtests were both adapted directly from the tests developed for use by the Army. Five of the ten items on the original Comprehension subtest were taken directly, some with only slight modifications, from the Comprehension test of the Army Individual Examination, while at least three other items were taken from the Army Alpha (Yerkes 1921). Items on the Arithmetic subtest were also adapted from the Army Individual Examination. Digit Span was derived from the Binet scales. Although Wechsler (1944) described in detail how he developed the word list for his optional vocabulary test, it too was derived from the Stanford-Binet test that was used in the Army.

Wechsler's Block Design was derived from a test constructed by Samuel Kohs called Kohs Block test (Kohs 1920). The Army Performance Scale had a somewhat similar test called Cube Construction. Wechsler (1944) notes that he changed the colors of the Kohs Block test, which had red, blue, yellow, and white blocks, to a simpler version using just two colors, red and white. This change might have been inspired by the Army's Cube Construction test in which the blocks had only two colors, red and natural wood (Yerkes 1921). Picture Arrangement also appeared on the individually administered

Army Performance Scale, although it had originally been developed for the Army Beta test. Wechsler adapted its pictures for his own test. Both Picture Completion and Digit Symbol were also borrowed from the Army Beta test. Digit Symbol (Coding) had been created by Otis and was based upon several versions of an older test called the substitution test, including the one designed by Woodworth (Yerkes 1921). Wechsler changed the administration time of Digit Symbol from 120 s to 90 s, although for the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III), the allotted time was restored to two minutes (Wechsler 1997). Object Assembly was derived from the Army Performance Scale.

Wechsler believed that his tests were primarily to be used as clinical instruments, despite the fact that they could also provide important quantitative information. Part of Wechsler's success and the popularity of his instruments can be attributed to his ability to look forward and remain ahead of the curve by developing tests that met the demands of the professionals that used them. However, more important than the tests he developed is the way in which Wechsler transformed the "field of IQ interpretation from *psychometric measurement* to *clinical assessment*" (Kaufman 2009, p. 37). Wechsler's view of intelligence testing was that it was only one part of personality testing. He believed that an individual evaluation was not complete without an assessment of non-intellective factors such as temperament, morality, and social values (Edwards 1974; Wechsler 1943, 1950). He included items in his scales that would give the examinee an opportunity to express rational thinking, purposeful thought, and effective problem solving. For example, because he felt it was a clinically useful test for diagnosing low mental ability, Wechsler included Digit Span in the original Wechsler-Bellevue even though it was a poor measure of *overall* intelligence (Wechsler 1944). According to Kaufman and Kaufman (2001):

> Wechsler embraced the inclusion of items with clinical content in his test batteries, believing that they enhanced the more complete measurement of intelligent behavior and improved the value of the psychometric instrument as a clinical tool. Subsequent to Wechsler's death, the publisher of revisions of his batteries yielded to political correctness and removed virtually all of the clinically charged items, the very ones that Wechsler believed would assess aspects of EI [emotional intelligence]. (p. 260)

Theory of the Wechsler-Bellevue Test

Wechsler did not base his tests on any hierarchical theories of intelligence, although he clearly saw the subtests as measuring different aspects of overall general intelligence or *g*. The most recent editions of the Wechsler scales (Wechsler 2008, 2012, 2014) have all been updated to more accurately reflect current psychological and neuropsychological research as well as some aspects of a three-stratum, hierarchical theory of intelligence known as the Cattell-Horn-Carroll (CHC) theory (Flanagan et al. 2013). Confirmatory actor analysis supports the fit of the Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV; Wechsler 2012), WAIS- IV, and WISC-V into this model (Lichtenberger and Kaufman 2013; Raidford and Coalson 2014; Wechsler 2014). However the original theory behind the Wechsler-Bellevue and subsequent Wechsler scales is based on the author's definition of intelligence (Wechsler 1939):

> The aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment. It is global because it characterizes the individual's behavior as a whole; it is aggregate because it is composed of elements or abilities which, though not entirely independent, are qualitatively differentiable. (p. 3)

According to Wasserman and Kaufman (in press), Wechsler believed that this definition served to encompass the various theories that had influenced him. So when he speaks of "The aggregate or global capacity," he could certainly have been referring to Spearman's general intelligence factor. His emphasis on intelligence being "composed of elements or abilities" is similar to Thurstone's primary mental abilities. When Wechsler refers to the capacity "to think rationally," was he acknowledging Terman's emphasis on intelligence being related to the capacity for abstract thinking? Finally Wechsler's definition includes the concept of dealing effectively with one's environment, which pays tribute to Binet's notion of intelligence being the ability to adapt to different circumstances.

The Controversial Aspects of Intelligence Testing

Wechsler was a staunch defender of the utility and necessity of proper intelligence testing, something he referred to as an "intelligent test" (Wechsler 1966). In a letter to the *New York Times* (Wechsler 1949b), he noted that IQ tests were reliable and valid and were more "definitely diagnostic" than an electrocardiogram. However, intelligence testing was controversial in Wechsler's day and remains controversial today (Kaufman 2009). Part of the reason for this controversy can be attributed to some of Wechsler's contemporaries who advocated the use of intelligence tests in support of their opinions on race, eugenics, and immigration. For example, in the first chapter of the Stanford-Binet manual published in 1916, Terman was explicit about his goals for the Stanford-Binet test. It was his hope that the routine testing of "feebleminded" individuals would result in some type of government action to limit their ability to reproduce. For Terman, the primary use of the intelligence test was to identify "mental defectives." The link between crime, vice, "industrial inefficiency," and people with limited mental abilities was unequivocal.

Goddard advocated for the use of mental tests to identify the feebleminded among prospective immigrants to the United States. In 1912, Goddard visited Ellis Island and administered the Binet test to an immigrant that he perceived as being feebleminded. The immigrant spoke no English, and the test was administered with the help of an interpreter. The results of the test confirmed Goddard's suspicions: the immigrant was mentally deficient. Goddard remained unswayed even when the interpreter told him that upon his own arrival to the United States, he himself would have been unable to answer many of the items on the test (Schultz and Schultz 2012).

The Army test data that had been gathered during the war were made available to psychologist Carl Brigham, an assistant professor at Princeton University. In 1923, he published *A Study of American Intelligence*, summarizing the data of the various Army tests. The foreword to the book, written by Yerkes, warned that the United States could not "afford to ignore the menace of race deterioration or the evident relations of immigration to national progress and welfare" (Brigham 1923). The data showed that the admittance of "intellectually inferior" immigrants, defined as those immigrants from non-Nordic countries, contributed to the decline of American intelligence. According to Brigham, the solution was to heavily restrict immigration to the United States, and in 1924 the US Congress did enact legislation that sharply curtailed levels of immigration (Kamin 1974).

It is worthwhile to correct a misperception that has recently appeared in the literature regarding the lower scores of non-native English speakers on the Army Beta. Ortiz et al. (2012) wrote:

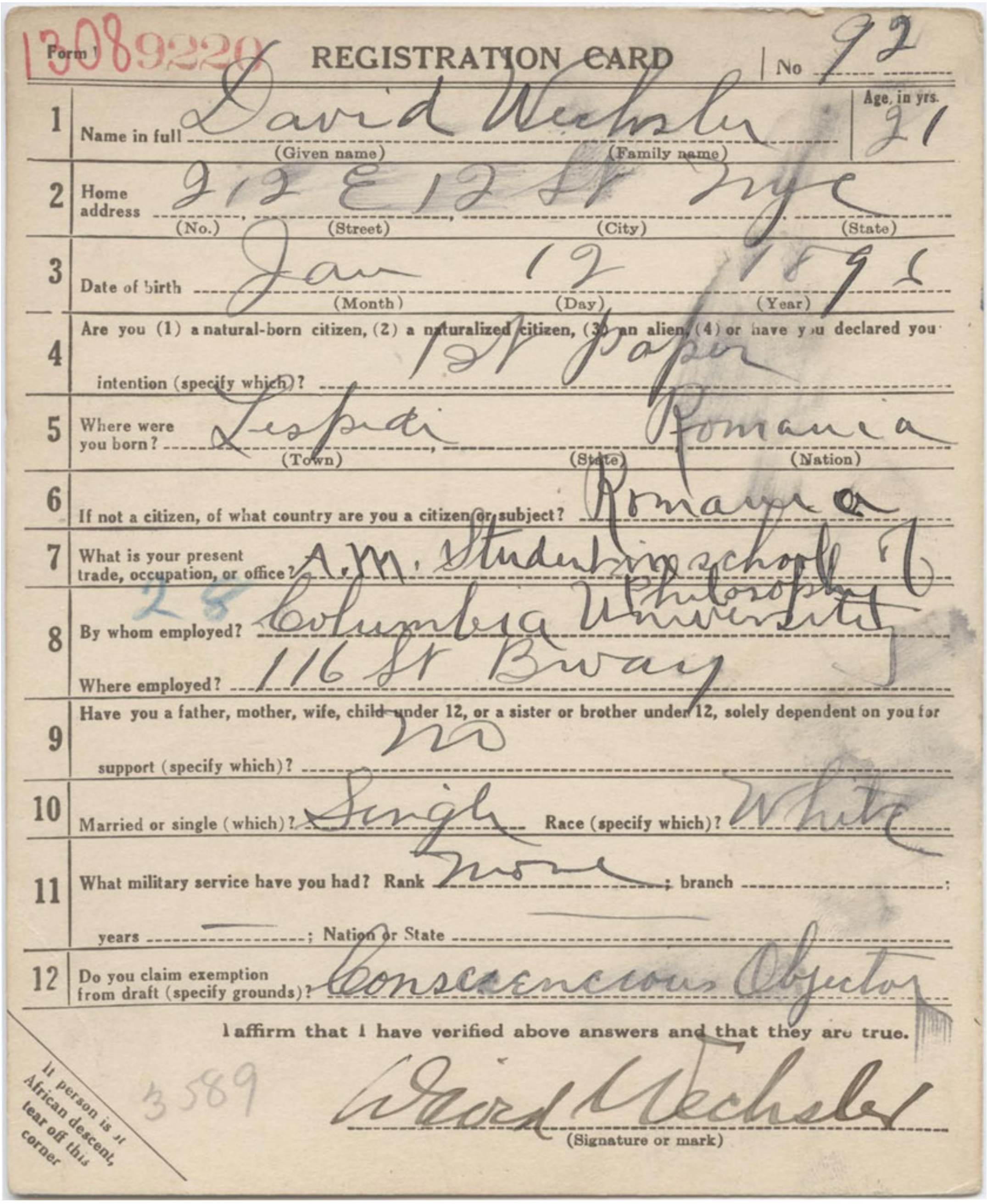
> In examining his data on nearly 1.75 million American men, Yerkes noted that the average raw score on the Beta for native English speakers was a stout 101.6, which classified them as Very Superior (Grade A). In contrast, the average raw score for non-native English speakers (who could also not read English) came in at only 77.8, which classified these individuals as Average (Grade C). For Yerkes, as well as the contingent of other notable psychologists working with him (e.g., Carl Brigham, David Wechsler, Lewis Terman), the results confirmed their own beliefs—that immigrants, particularly those from certain countries and from lower classes, were merely displaying their inherited lack of intelligence. (pp. 537–538)

Ortiz et al. are correct in noting the manner in which some of psychologists misinterpreted the Army data as a result of their prejudiced views regarding immigrants. There is no evidence to suggest that Wechsler had the same beliefs regarding the intelligence of immigrants as did Yerkes, Brigham, and Terman. He was a junior member of psychology staff in the Army and had yet to become a notable psychologist. In fact, as an Eastern European immigrant, persecuted because of his religion, Wechsler would likely have disagreed with the beliefs of his older contemporaries. Wechsler developed his test in part as an attempt to correct some of the flaws that existed with earlier tests, such as the emphasis on verbal tests. He defended intelligence tests against accusations of cultural and racial bias. As mentioned previously, Wechsler believed that on the whole people are not very different from each other (Wechsler, 1935b). Furthermore, Wechsler was careful to note that he had not included a sample of African-Americans when he developed the Wechsler-Bellevue and cautioned against using the test with that population. (Wechsler 1944).

In 2011, Fox et al. published the results of a study from the Bucharest Early Intervention Project (BEIP). The goal of this project was to introduce early intervention services and foster care into Romania and to study the results over a period of a decade. The authors examined the IQ scores of 103 participant in the BEIP, eight years after they entered the program. All of the children participating in the study had come from deprived home settings. The researchers found that the children who had participated in the early intervention program had significantly higher scores than a control group of deprived children who had not participated in the program. In their methods section, the authors report that the participants' IQs were measured using the WISC-IV (2003). It seems almost poetical that over a century after he was forced to leave Romania, his native land, one of Wechsler's tests was used as part of a program there to improve the quality of life of deprived children. In light of the results achieved by Fox et al. (2011), Wechsler's (1966) words in the *New York Times Magazine* ("The I.Q. is an Intelligent Test") provide yet another example of his forward-thinking approach to intelligence testing:

> It is true that the results of intelligence tests, and of others, too, are unfair to the disadvantaged, deprived . . . but it is not the I.Q. that had made them so. The culprits are poor housing, broken homes, a lack of basic opportunities, etc. etc. If the various pressure groups succeed in eliminating these problems, the I.Q.'s of the disadvantaged will take care of themselves.

Appendix


13089220 Form 1 REGISTRATION CARD No 92

1	Name in full David Wechsler (Given name) (Family name)	Age, in yrs. 21
2	Home address 312 E 12 St Nyc (No.) (Street) (City) (State)	
3	Date of birth Jan 12 1896 (Month) (Day) (Year)	
4	Are you (1) a natural-born citizen, (2) a naturalized citizen, (3) an alien, (4) or have you declared your intention (specify which)? 1st paper	
5	Where were you born? Lespedi Romania (Town) (State) (Nation)	
6	If not a citizen, of what country are you a citizen or subject? Romania	
7	What is your present trade, occupation, or office? A.M. Student in school of Philosophy	
8	By whom employed? Columbia University. Where employed? 116 St Bway	
9	Have you a father, mother, wife, child under 12, or a sister or brother under 12, solely dependent on you for support (specify which)? No	
10	Married or single (which)? Single Race (specify which)? White	
11	What military service have you had? Rank None; branch ; years ; Nation or State	
12	Do you claim exemption from draft (specify grounds)? Conscientious Objector	

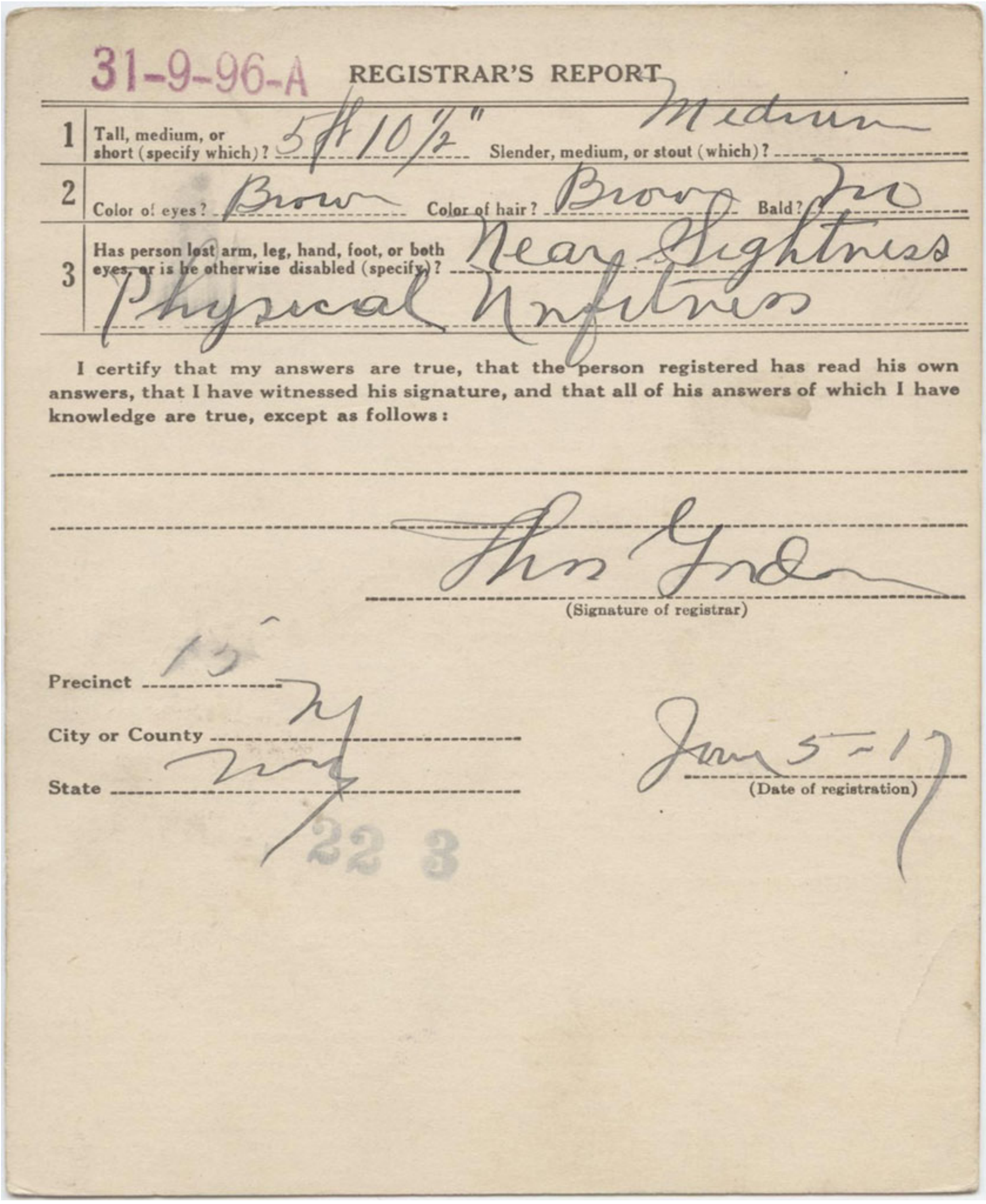
I affirm that I have verified above answers and that they are true.

David Wechsler (Signature or mark)

If person is of African descent, tear off this corner

3589

Fig. 12.1 The front and back of David Wechsler's World War 1 registration card. On June 5, 1917, all men ages 21–31 in the United States were required to register to determine eligibility for induction into the military. The response to question 4 "1st paper" indicates that Wechsler was eligible for the draft. It is likely that the A.M. in the response to question 7 stands for *artium magister*, or master of arts. From the National Archives and Records Administration. In the public domain



31-9-96-A REGISTRAR'S REPORT

1 Tall, medium, or short (specify which)? 5 ft 10½" Slender, medium, or stout (which)? Medium

2 Color of eyes? Brown Color of hair? Brown Bald? No

3 Has person lost arm, leg, hand, foot, or both eyes, or is he otherwise disabled (specify)? Near Sightness Physical Unfitness

I certify that my answers are true, that the person registered has read his own answers, that I have witnessed his signature, and that all of his answers of which I have knowledge are true, except as follows:

(Signature of registrar)

Precinct 15

City or County NY

State NY

Jun 5 - 17
(Date of registration)

22 3

Fig. 12.1 (continued)

Fig. 12.2 View of Dr. David Wechsler seated at a table, conducting a test on a patient. The date of the photo is unknown. The machine might be a version of the photo-galvanograph that Wechsler patented in 1924 (Image courtesy of The Lillian and Clarence de la Chapelle Medical Archives at NYU)

References

Becker, K. A. (2003). *History of the Stanford-Binet intelligence scales: Content and psychometrics.* (Stanford-Binet intelligence scales, fifth edition, Assessment Service Bulletin No. 1). Itasca: Riverside. Retrieved from http://www.riverpub.com/products/sb5/pdf/SB5_ASB_1.pdf

Benson, N., Hulac, D. M., & Kranzler, J. H. (2010). Independent examination of the Wechsler adult intelligence scale—Fourth edition (WAIS-IV): What does the WAIS-IV measure? *Psychological Assessment, 22*(1), 121–130.

Binet, A., & Simon, T. (1916a). New methods for the diagnosis of the intellectual level of subnormals. In H. H. Goddard (Ed.), *The development of intelligence in children (the Binet-Simon Scale)* (trans: Kite, E. S.; pp. 37–90). Baltimore: Williams & Wilkins. (Original work published 1905).

Binet, A., & Simon, T. (1916b). Upon the necessity of establishing a scientific diagnosis of inferior states of intelligence. In H. H. Goddard (Ed.), *The development of intelligence in children (the Binet-Simon Scale)* (trans: Kite, E. S.; pp. 9–36). Baltimore: Williams & Wilkins. (Original work published 1905).

Boake, C. (2002). From the Binet–Simon to the Wechsler–Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology, 24*, 383–405.

Brigham, C. C. (1923). *A study of American intelligence*. Princeton: Princeton University Press.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15*, 373–381.

Cattell, J. M., & Farrand, L. (1896). Physical and mental measurements of the students of Columbia University. *Psychological Review, 3*(6), 618–648. doi:10.1037/h0070786.

Cornbise, A. E. (1997). *Soldiers-scholars: Higher education in the AEF, 1917–1919*. Philadelphia: American Philosophical Society.

DeSipio, L. (1987). Social science literature and the naturalization process. *International Migration Review, 21*(2), 390–405.

Edwards, A. J. (1974). *Selected papers of David Wechsler*. New York: Academic.

Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (2nd ed.). Hoboken: Wiley.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. (2013). *Essentials of cross-battery assessment* (3rd ed.). Hoboken: Wiley.

Fox, N. A., Almas, A. N., Degnan, K. A., Nelson, C. A., & Zeanah, C. H. (2011). The effects of severe psychosocial deprivation and foster care intervention on cognitive development at 8 years of age: Findings from the Bucharest Early Intervention Project. *Journal of Child Psychology and Psychiatry, 52*(9), 919–928. doi:10.1111/j.1469-7610.2010.02355.x.

Georgas, J., Weiss, L. G., van de Vijver, F. J. R., & Saklofske, D. H. (2003). *Culture and children's intelligence: Cross-cultural analysis of the WISC–III*. San Diego: Academic Press.

Goddard, H. H. (Ed.). (1916). *The development of intelligence in children (the Binet-Simon Scale)* (trans: Kite, E. S.). Baltimore: Williams & Wilkins.

Guilford, J. P. (1972). Thurstone's primary mental abilities and structure-of-intellect abilities. *Psychological Bulletin, 77*(2), 129–143. doi:10.1037/h0032227.

Iordachi, C. (2002). Citizenship and national identity in Romania: A historical overview. *Regio: Minorities, Politics, Society: English Edition, 1*, 3–34.

Jones, H. E., & Wechsler, D. (1928). Galvanometric technique in studies of association. *American Journal of Psychology, 40*(4), 607–612.

Kamin, L. J. (1974). *The science and politics of I.Q.* Potomac: Erlbaum.

Kaufman, A. S. (2009). *IQ testing 101*. New York: Springer.

Kaufman, A. S., & Kaufman, J. C. (2001). Emotional intelligence as an aspect of general intelligence: What would David Wechsler say? *Emotion, 1*(3), 258–264. doi:10.1037/1528-3542.1.3.258.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler intelligence scale for children—Fourth edition: What does it measure? *School Psychology Review, 35*, 108–127.

Kohs, S. C. (1920). The block-design tests. *Journal of Experimental Psychology, 3*(5), 357–376. doi:10.1037/h0074466.

Lichtenberger, E. O., & Kaufman, A. S. (2013). *Essentials of WAIS-IV assessment* (2nd ed.). New York: Wiley.

Louttit, C. M., & Browne, C. G. (1947). The use of psychometric instruments in psychological clinics. *Journal of Consulting Psychology, 11*, 49–54.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). New York: Oxford University Press.

Matikainen, S. (2006). *Great Britain, British Jews and the international protection of Romanian Jews, 1900–1914: A study of Jewish diplomacy and minority rights*. Jyväskylä: University of Jyväskylä.

McNemar, Q. (1942). *The revision of the Stanford-Binet scale: An analysis of the standardization data*. Oxford: Houghton Mifflin.

Ortiz, S., Ochoa, S. M., & Dynda, A. M. (2012). Testing with culturally diverse and linguistically diverse populations: Moving beyond the verbal-performance dichotomy into evidence-based practice. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 526–552). New York: The Guilford Press.

Pillsbury, W. B. (1947). *Biographical memoir of James McKeen Cattell 1860–1944*. Washington, DC: National Academy of the Sciences.

Raiford, S. E., & Coalson, D. L. (2014). *Essentials of WPPSI-IV assessment*. Hoboken: Wiley.

Schultz, D. P., & Schultz, S. E. (2012). *A history of modern psychology* (10th ed.). Belmont: Wadsworth.

Spearman, C. (1904). "General Intelligence," objectively determined and measured. *The American Journal of Psychology, 15*(2), 201–292.

Stern, W. (1912). *Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung an Schulkindern*. [The psychological methods of testing intelligence and their application to school children.] Leipzig: J. A. Barth. Retrieved from http://openlibrary.org/books/OL24489667M/Die_psychologischen_Methoden_der_Intelligenzprüfung_und_deren_Anwendung_an_Schulkindern

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Terman, L. M. (1921). Intelligence and its measurement: A symposium–II. *Journal of Educational Psychology, 12*(3), 127–133. doi:10.1037/h0064940.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence*. Boston: Houghton-Mifflin. (Forms L and M of the Stanford-Binet intelligence scale).

Thurstone, L. L. (1938). *Primary mental abilities* (Psychometric monographs, Vol. 1). Chicago: University of Chicago Press.

U.S. Office of the Provost Marshal General. (1918). *Selective Service Regulations Prescribed by the President under the Authority Vested in Him by the Terms of the Selective Service Law (Act of Congress Approved May 18, 1917, with Supplementary and Amendatory Acts and Resolutions)* (2nd ed.). Washington, DC: GPO.

Wahlstrom, D., Breaux, K. C., Zhu, J., & Weiss, L. G. (2012). The Wechsler preschool and primary scale of intelligence—Third edition, the Wechsler intelligence scale for children—Fourth edition, and the Wechsler individual achievement test—Third edition. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 224–248). New York: The Guilford Press.

Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 3–55). New York: The Guilford Press.

Wasserman, J. D., & Kaufman, A. S. (in press). Biography of David Wechsler. In R. L. Cautin & S. O. Lilienfeld (Eds.), *Encyclopedia of clinical psychology*. Malden: Wiley Blackwell.

Wechsler, D. (1917a). A study of retention in Korsakoff psychosis. *Psychiatric Bulletin, 2*, 403–451.

Wechsler, D. (1917b). *World War One Selective Service System draft registration cards, 1917–1918* (Administration. M1509, Roll: 1765679, Draft Board: 96). Washington, DC: National Archives and Records Administration.

Wechsler, D. (1918). *Petition for naturalization, David Wechsler. United States District Court of the Northern District of Georgia, Rome Division*. Morrow: National Archives and Records Administration.

Wechsler, D. (1924). U.S. Patent No. 1,503,401. Washington, DC: U.S. Patent and Trademark Office.

Wechsler, D. (1932). Analytic use of the Army Alpha examination. *Journal of Applied Psychology, 16*(3), 254–256. doi:10.1037/h0072688.

Wechsler, D. (1935a). The concept of mental deficiency in theory and practice. *Psychiatric Quarterly, 9*, 232–236.

Wechsler, D. (1935b). *The range of human capacities*. Baltimore: The Williams and Wilkins Company.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams and Wilkins.

Wechsler, D. (1943). Nonintellective factors in general intelligence. *Journal of Abnormal and Social Psychology, 38*, 101–103.

Wechsler, D. (1944). *The measurement of adult intelligence* (3rd ed.). Baltimore: The Williams and Wilkins Company.

Wechsler, D. (1946). *The Wechsler-Bellevue intelligence scale, form II. Manual for administering and scoring the test*. (WB-II). New York: The Psychological Corporation.

Wechsler, D. (1949a). *Wechsler intelligence scale for children (WISC)*. New York: The Psychological Corporation.

Wechsler, D. (1949b). Precision instrument [Letter to the editor], *New York Times Magazine*, p. SM4. Retrieved from http://query.nytimes.com/mem/archive/pdf?res=FB0F17F73958107A93C0A9178BD95F4D8485F9

Wechsler, D. (1950). Cognitive, conative, and non-intellective intelligence. *American Psychologist, 5*, 78–83.

Wechsler, D. (1952). *The range of human capacities* (2nd ed.). Baltimore: The Williams and Wilkins Company.

Wechsler, D. (1966). The I.Q. is an intelligent test. *New York Times Magazine*. p. 191. Retrieved from http://query.nytimes.com/mem/archive/pdf?res=F20A15FE3F55117B93C4AB178DD85F428685F9

Wechsler, D. (1997). *Wechsler adult intelligence scale—Third edition (WAIS-III)*. San Antonio: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler intelligence scale for children-fourth edition (WISC-IV)*. San Antonio: The Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale -fourth edition (WAIS-IV)*. San Antonio: Pearson.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence-fourth edition (WPPSI-IV)*. San Antonio: Pearson.

Wechsler, D. (2014). *Wechsler intelligence scale for children-fifth edition: Technical and interpretive manual*. Bloomington: Pearson.

Wechsler, D., & Naglieri, J. A. (2006). *Wechsler nonverbal scale of ability (WNV)*. San Antonio: Pearson.

Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., & Maerlender, A. (2004). *Wechsler intelligence scale for children-fourth edition-integrated (WISC-IV Integrated)*. San Antonio: The Psychological Corporation.

Winston, A. S. (1996). "As his name indicates": R.S. Woodworth's letters of reference and employment for Jewish psychologists in the 1930s. *Journal of the History of the Behavioral Sciences, 32*(1), 30–43.

Woodworth, R. S. (1918). *Dynamic psychology, by Robert Sessions Woodworth*. New York: Columbia University Press.

Woodworth, R. S., & Wells, F. (1911). Association tests. *The Psychological Monographs*, *13*(5). doi:10.1037/h0093064.

Yerkes, R. M. (Ed.). (1921). *Memoirs of the National Academy of Sciences: Vol. 15. Psychological examining in the United States Army*. Washington, DC: US Government Printing Office.

Yerkes, R. M., Bridges, J. W., & Hardwick, R. S. (1915). *A point scale for measuring mental ability*. Baltimore: Warwick & York.

A.R. Luria and Intelligence Defined as a Neuropsychological Construct

13

Dana Princiotta and Sam Goldstein

Gedenkschrift to the Father of Neuropsychology

July 16, 2002, marked an important date in the world of neuropsychology. This centennial day celebrated the significant contributions of the Russian neuropsychologist Alexander Romanovich Luria, born in 1902. A world-renowned scientist, Luria's theories continue to excite experts more than 30 years after his death. By the 1980s, following Luria's death (August 14, 1977), a survey of neuropsychologists found that he was revered as "number 1" among the ten founders of neuropsychology (Akhutina and Pylaeve 2011). Underestimating this legacy in a single chapter is an unfortunate inevitability when asked to chronicle Luria's unequivocal impact on the practice of psychology and neuropsychology worldwide.

Procuring the title "father of neuropsychology" necessitated contributions in both the psychological and medical realms as neuropsychology was growing to be a recognized field of study. One of Luria's many acclaimed works, *The Working Brain* (1973a), was written by only 30 years into clinical neuropsychology's gestation (Cole et al. 2006). The immense magnitude of Luria's legacy is not fully appreciated by most scholars because of the political and linguistic deterrents surrounding Luria's work. Despite publishing extensively over the course of 50 years, a complete biography encompassing all published works does not exist (Cole et al. 2006). Rationales for this are associated with challenges in physically locating his publications and in translating works written in Russian. In Luria's 50-year contribution to neuropsychology, he authored more than a lifetime's worth of work. Nearly 40 years following his death, we continue to reflect on Luria's philosophies.

Born to Jewish parents in Kazan Russia, Luria's parents pursued careers in medicine and dentistry despite political tensions in Russia. Luria's father (Roman Albertovich Luria) was a physician at the University of Kazan; his mother was a dentist. Luria's sister followed the familial calling and pursued psychiatry. Luria's formal pursuit of higher education began at the age of 16 when he was accepted at Kazan State University. Against his father's wishes, Luria pursued psychology. During his studies at the University, he started the Kazan Psychoanalytic Association. He graduated in 1921 at the age of 19. Luria completed his studies while the Russian Revolution was underway. In 1924, Luria was introduced to

D. Princiotta, Ph.D. (✉)
1845 West Orange Grove Road,
Tucson, AZ 85704, USA
e-mail: danakprinciotta@gmail.com

S. Goldstein, Ph.D.
University of Utah, 230 South 500 East, Suite 100,
Salt Lake City, UT 84102, USA
e-mail: info@samgoldstein.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_13,

Lev Semionovich Vygotsky, setting the stage for a lifetime collaboration and the inauguration of Luria's 50-year career.

The Vygotsky Circle

Much of Luria's earlier documented work arose under the guidance of Vygotsky. Luria, Vygotsky, and a team of scholars including medical specialists, neuroscientists, psychologists, and educational specialists collaborated to investigate the processes and psychology involved in an individual's maturation as it related to culture (Luria, Cole, & Cole, 1979; Cole 2005a). This collection of scholars would become known as the Vygotsky Circle (Luria, Cole, & Cole, 1979). Exploiting the terms *historical, cultural*, and *instrumental*, the circle aspired to "discover the way natural processes such as physical maturation and sensory mechanisms become intertwined with culturally determined processes to produce the psychological functions of adults" (Luria, Cole, & Cole, 1979, p. 43). Much of this work was facilitated in the early years by the Institute of Psychology in Moscow.

The term "cultural-historical psychology" was one of the prominent earliest contributions from the collective Vygotsky Circle. Ideas, theories, and research flourished under Luria, Vygotsky, and Alexey Leontiev's guidance (Cole 2005a). The Vygotsky Circle paid tribute to the interaction of nature/nurture and natural/cultural factors contributing to the development of the mind, with particular focus on the cultural, historical, and instrumental aspects of psychology (Kotik-Friedgut 2006). They aimed to "discover the way natural processes such as physical maturation and sensory mechanisms become intertwined with culturally determined processes to produce the psychological functions of adults" (Luria, Cole, & Cole, 1979, p. 43).

Within a cultural-historical framework, the underpinnings of Vygotsky's work supported Luria in speculating that cognitive processes in humans are organized in a way that thought and language are closely related, suggesting that spoken language reflects thoughts (Kaczmarek 1999). Luria expounded further and proposed that children are first introduced to linguistic commands by adults and then in turn adopt "self-commands" beginning as commands spoken aloud and then internally; thus, language mediates human behaviors (Kaczmarek 1999). "The development of any type of complex conscious activity is at first expanded in character and requires a number of essential aids for its performance; not until later does it gradually become condensed and converted into an automatic (motor) skill" (Luria 1973a; Kotik-Friedgut 2006, p. 48).

Luria examined the developmental differences as they relate to nature and nurture with his famous twin studies in the 1930s. He studied identical and fraternal twins at the Medico-Genetic Institute in Moscow. Luria believed that genetic contributions to behavior manifested expressly during tasks requiring natural cognitive processes rather than tasks relying on "culturally mediated processes" (Cole et al. 2006, p. 83). Working with five sets of identical twins, Luria utilized differing levels of instructions for twin dyads. The results suggested almost equal abilities in the tasks requested by the examiners; however, differences were apparent in their language and the ability to repeat the task from memory 18 months later (Bowden 1971). This implied that functions could be considered acquired, as the genetic factor was held constant in this study (Bowden 1971).

Subsequent studies focused on the development of speech in 5-year-olds with minimal speech development. A twin that was encouraged to verbalize his/her thoughts and desires developed speech much quicker than a twin confined to an "indifferent" adult, thereby highlighting the importance of the social component necessary to acquire language (Bowden 1971). This was best summarized by Kotik-Friedgut in 2006: "We always speak about things we see. We never speak about things we did not see. Prior to Luria's way of thought, the world was confined to concrete thinking; after Luria had his say, we were all more of the thinking that the ability to think abstractly was due to schooling and culture" (Kotik-Friedgut 2006, p. 44).

World War II and the Emergence of Neuropsychology

In the late 1930s, the Vygotsky Circle dispersed as German soldiers invaded the Soviet Union. The influence of the Vygotsky Circle resurfaced following the end of World War II. Due to mounting tensions in Soviet Russia in the 1930s, Luria embarked on a degree from medical school and focused much of his attention toward the aphasias. It is not surprising given Luria's intense interest in the acquisition of language and the social environment as demonstrated with the Vygotsky Circle that Luria would turn his attention toward the aphasias. Luria was fascinated with the aphasias, particularly in the classification of sounds, speech production and in writing associated with aphasias (Kolb and Whishaw 2009).

The aftermath of World War II yielded numerous case studies by Luria of individuals suffering from various forms of aphasia (Bowden 1971). For example, Luria discovered that those individuals suffering from oral apraxias could manipulate speech sounds with the utilization of observing their own movements of the lips and tongue in a mirror, thus utilizing an "environmental component of the reorganized functional system" (Bowden 1971, p. 413). Those who study Luria's work closely recognize his association with traumatic aphasia. Luria devoted a book on this topic, outlining case studies of hundreds of individuals suffering from localized brain injuries during World War II (Cole et al. 2006). "It was during the war and its aftermath that neuropsychology became a full-fledged science" (Cole et al. 2006, p. 140).

One such case was *The Man with a Shattered World*. A World War II veteran named Saletsky suffered from brain damage from a bullet entering the posterior left hemisphere intersections of the parietal, occipital, and temporal cortexes (Kolb and Whishaw 2009). Luria would come to follow Saletsky's trials and tribulations and progress for 26 years. Saletsky never returned to a premorbid level of functioning (Kolb and Whishaw 2009). Stated by Luria, "the damaged areas of the cerebral cortex could not be restored. Hence when he tried to think, his mind had to detour around these scorched areas and employ other faculties with which to learn and try to recover some lost skills" (Luria 1972, p. 158; Kolb and Whishaw 2009). This period became the breeding ground for Luria's systematic approach to the brain and mental functions forming the backbone of neuropsychology (Cole 2005a).

In 1959, Luria described a second case study focused on a war wound suffered by V., an officer of the Polish army. V. was wounded in the left occipital-parietal region during the war (Luria 1959). Although speech and vision improved and V. recovered in these areas after 6 months, V. complained of "attacks of giddiness and nausea as well as photophobia and epiphora" (Luria 1959, p. 439). V. demonstrated defects of visual perception and oculomotor ataxia, among other difficulties. World War II yielded a forum in which medical and psychological scholars would come to study the brain more intimately than before, especially as it relates to traumatic brain injuries. "To signify the combination of these two enterprises, the 'neurological' and 'psychological,' the term neuropsychology was coined" (Cole et al. 2006, pp. 157, 158).

Following the war, Luria's work was interrupted when he was asked to leave the Institute of Neurosurgery due to his Jewish background. Eventually, he was invited to return. He continued his study of neuropsychology in the 1950s at the Institute of Neurosurgery until his time of death (Cole 2005a). For the field of neuropsychology to advance, Luria stated that neuropsychologists needed to focus on a more thorough understanding of the "neurophysiological mechanisms" that were responsible for operating brain structures (Cole 2005a).

Topography, Localization, and Functioning of the Brain

In the early half of the nineteenth century, phrenology, a then popular philosophy developed by a German physician, claimed that certain brain areas had localized and specific functions (Kolb and Whishaw 2009). Under the guidance of phrenology, Franz Joseph Gall concluded that the size of the brain areas was a significant indicator as to human behaviors (Kolb and Whishaw 2009). Now considered pseudoscience, neurologists and physiologists in the latter part of the nineteenth century continued to develop their understanding of the brain.

Although neurologists and psychiatrists of the 1880s were able to develop functional maps of the cerebral cortex (Cole et al. 2006), our understanding was greatly enhanced half a century later in the 1930s with Luria's assistance. "The discovery that a complex form of mental activity can be regarded as the function of a local brain area aroused unprecedented enthusiasm in neurological science. Within a short time many other brain centers for intellectual functions were found…" (Cole et al. 2006, p. 210). Our current conceptualization of the topography of the brain as well as localization of function was made possible by Luria's significant contribution.

Luria elaborated, "To say that the human brain operates as a whole means to make simultaneously a correct and an erroneous statement. It is correct because the most complex forms of human actions require the participation of all brain systems; it is erroneous because we can hardly admit that the Human Brain—this highest point of Evolution—works as an undifferentiated whole and that the quality of its work depends exclusively on the active mass of its excited tissue" (Luria 1969b, p. 9). The complexity of this statement would fuel much of Luria's work in the 1930s and beyond.

Multiple concepts existed in the 1950s related to brain functioning—the most popular entailed matching a specific mental function with a specific cortical area or matching all brain functions to all behaviors (Cole et al. 2006). Luria preferred to focus on a complete functional system rather than an isolated function (Luria 1973a). A complete functional system allows for the conceptualization of multiple components working together. "This means that all these apparently so widely different functions incorporate a common factor, and it allows an approach to be made to the more intimate analysis of the structure of psychological processes" (Luria 1973b, p. 42).

One of the major research questions at that time was whether a complex functional system could be localized. Luria did not suggest that this complex system could be grouped and located in zones of the cortex either (Luria 1973a). Rather, he suggested that different areas of the brain must be involved in this system and not necessarily neighboring each other nor even located near one another (Luria 1973a). At this point in time, two major camps were developing revolving around Luria's work, those favoring a localized view of the brain and those favoring an integrated system.

The terms "function" and "localization" have come under great scrutiny as a result of Luria's advancements in the field of neuropsychology. While neurologists may argue one point about function and localization, a neuropsychologist may argue another. In the words of Luria himself: "Supporters of the first (narrow localization) approach started with the viewpoint that both elementary and higher (mental) functions must be viewed as an immediate function of narrowly limited parts of the brain; therefore they found it possible to speak of zones in which such phenomena as motor and sensory images of words, the function of writing or counting, are *localized*, and considered that the loss of these functions is an unequivocal symptom of damage in a corresponding zone of the brain cortex. Supporters of the second (anti-localization) approach, outwardly beginning from the opposite conception, in fact share the principal position of their opponents" (Luria 1964, p. 5).

In further illustrating his tenets of localization and functioning, Luria equated the meaning of the term function with that of bodily organs, including the function of the liver or pancreas, further arguing that the function of a particular brain area cannot be likened to the functioning of an organ (Luria 1973a). To think that most mental processes could be using "isolated or even indivisible faculties, which can be presumed to be the direct function of limited cell groups or to be localized in particular brain areas" cannot be true (Luria 1973b, p. 29). "Furthermore, these functional systems are not in 'narrow zones of the cortex' but must be located throughout the brain, regardless of location from another team member of the functional system" (Luria 1973b, p. 31).

Luria believed that one could not reduce this question to a simple concept in which a disturbance of a mental function is the direct result of the destruction of a specific part of the brain—thus providing confirmation that the function is localized in the damaged portion of the brain (Luria 1973b, p. 34). "It should be apparent that if the operation of intellectual processes is thought of in terms of functional systems instead of discrete abilities, we have to reorient our ideas about the possibility of localizing intellectual functions … our solution has been to think of the functional system as a working constellation of activities with a corresponding working constellation of zones of the brain that support the activities" (Cole et al. 2006, p. 141). This might be one of the biggest distinguishing features between the human and animal brain (Luria 1973b, p. 31). The meaning of localized brain damage as it relates to these higher mental functions began for Luria through the Vygotsky Circle (Cole et al. 2006).

Luria did not make the mistake of attempting to map out the precise locations with specific higher cognitions occurring when he explained his three functional units of the brain. To do so would suggest that parts of the brain were functioning independently. In other words, one could not select a cognitive task and assign the task as only relying on one type of processing or ability (Luria 1973a; S. Goldstein, personal communication, July 13, 2013). In Luria's words, "…perception of memorizing gnosis, and praxis, speech and thinking, writing, reading and arithmetic cannot be regarded as isolated or even indivisible faculties…" (Luria 1973b, p. 29).

Luria is perhaps most remembered for his teachings in the organization of functional brain systems including energizing, coding, and planning and the cultural contribution of the environment (Cole 2005a)—the key point aimed at the methodological conjunction of both theory and practice (Cole 2005a, p. 40). The essential tenets of the functional structure and brain organization, or higher mental functions, began with Vygotsky and continued with Luria (Akhutina and Pylaeve 2011). Rather than taking a "horizontal" viewpoint, Luria focused our attention to "vertical" perspectives of both the surface brain structures and deeper brain structures (Cole et al. 2006, p. 159). A discovery that the brain structures responsible for cortical tone are actually housed in the subcortex and brain stem (i.e., the structures influence tone and are regulated by the structures) radically altered the trajectory of neuropsychology (Luria 1973a).

Higher Mental Functions

Over the course of his impressive career, Luria was engaged in diverse case studies of individuals with extraordinary abilities and deficits alike. While we have briefly touched upon two of Luria's famous case studies demarcating deficits, the case of S. chronicles a case of greater faculties. In the *Mind of a Mnemonist* (1968), Luria described one of his most famous case examples, S., a man with fascinating abilities in the area of memory. S. first came to Luria after S.' employer noted that he never took notes during meetings (Kolb and Whishaw 2009). Remarkably, S. was able to visualize stimuli mentally and recall stimuli by reading from an internal "photocopy" of the original (Kolb and Whishaw 2009). Interestingly, S. also met diagnostic criteria for synesthesia, an ability to perceive a stimulus of one sense as the sensation of a different sense (e.g., tasting colors) (Kolb and Whishaw 2009). What allowed S. to create an internal photocopy of the original? How do we begin to understand higher mental functions like S.' extraordinary memory?

In 1966, Luria described stages of functions that were imperative for the development of intelligence and executive functions (S. Goldstein, personal communication, July 13, 2013; Vygotsky, Veer, & Valsiner 1994). Only a few years later, the systemic-dynamic approach of brain organization of "higher mental functions" was refined through the work of the Vygotsky Circle (Kotik-Friedgut 2006). Luria described interconnected levels that assisted in explaining the relationship between the brain and behavior and neurocognitive disorders. These levels included the structures of the brain, the functional organization (based on structure), syndromes and impairments arising in brain disorders, and clinical methods of assessment (Kaczmarek

1999; S. Goldstein, personal communication, July 13, 2013).

Recall the cultural-historical theory in which Vygotsky and Luria proposed that environmental interactions were particularly salient in the brain and behavior. Expanding on Vygotsky's work, Luria suspected that higher mental functions (e.g., abstraction, memory, and attention), a term that will be repeated frequently in this chapter, have an origin in social matters (e.g., language and thought processes) and are "complex and hierarchical in structure" (Luria 1973b, p. 31; Kotik-Friedgut 2006; Goldstein, personal communication, July 13, 2013). This statement supports a theory in which human conscious activity is developed with the utilization of external aids or tools (Kotik-Friedgut 2006). Furthermore, "At the same time, the process of internalization in the development of higher mental functions takes place under the influence of a specific cultural context, thus shaping and moderating the process of development and the functioning of these basic cognitive abilities" (Kotik-Friedgut 2006, p. 43). As demonstrated in the case study of S., the optimal outcome is efficiency in cortical functioning in the areas of attention, memory, intelligence, executive function, and language.

Luria's particularly salient points are included in his books, *Higher Cortical Functions in Man* (1980) and *The Working Brain* (1973b). Luria (1973b, p. 43) proposed that the brain is comprised of three functional units, each of which are "hierarchical in structure and consist of at least three cortical zones built one above the other." In other words, each "unit" is further divided into three cortical zones, arranged vertically. Coined primary, secondary, and tertiary cortexes, Luria described the cortical areas as working together. As sensory information travels from the first to second to third zones, information in the tertiary zone may then be processed in the amygdala or the paralimbic cortex for memory and emotional processing (Kolb and Whishaw 2009). In *The Working Brain*, Luria described a complex functional system in which these three contributing units worked in constellation to "create" higher mental activities (Luria 1973a).

Providing further indication for these three units and their associated neuropsychological abilities and location in the brain, Luria stated that the brain stem, diencephalon, and medial regions of the cortex constitute the first functional unit (attention-arousal system) (Luria 1973a). More specifically, Luria was describing the midbrain, medulla, thalamus, and hypothalamus. Working together, these organs maintain appropriate cortical tone (Luria 1973a). "Recent formulations of these regions suggest that some structures at the level of the diencephalon and medial regions have reciprocal connections to the cortex through a variety of subcortical circuitries, potentially influencing a wide range of behaviors" (Koxiol 2009).

The occipital, parietal, and temporal lobes are associated with the second unit in the medial temporal portions positioned posterior to the central sulcus (Semrud-Clikeman, & Fine 2008; Goldstein, personal communication, July 13, 2013). The predominant role of the second unit is to process and retain external information such as sensory reception and integration. Here the sensory modality does correspond to a particular part of the brain (e.g., auditory stimuli to the temporal lobe) (Luria 1980). The second unit was under the guidance of three guidelines proposed by Luria: first, the makeup of cortical zones does not indeed remain the same during development; second, the specificity of cortical zone functioning decreases with development; and, third, with development, an increase in lateralization occurs (Luria 1980).

The third functioning unit is responsible for the regulation and evaluation of behavior, including self-monitoring (Luria 1980). The third functional unit, or the prefrontal area of the frontal lobes, "synthesizes the information about the outside world … and is the means whereby the behavior of the organism is regulated in conformity with the effect produced by its actions" (Luria 1980, p. 263). This unit manipulates the most complex components of human behavior such as consciousness, personality, voluntary activity, conscious impulse control, and linguistic skills (Das 1980; S. Goldstein, personal communication, July 13, 2013). Here we would expect to find the unit responsible for planning and executive functioning, namely, (McCloskey, Perkins, & Van Divner 2009; S. Goldstein, personal communication, July 13, 2013). Of the three units, the first and the third are the most related to one another; however, the second and the third unit rely on the first condition for collaboration efforts (Luria 1973a).

In other words, a person is permitted or able through these processes along with already acquired knowledge to navigate their fluid environments (Luria 1973a; S. Goldstein, personal communication, July 13, 2013).

Luria premised his hierarchical model on three main points:

1. Information is processed serially, one step at a time, by the brain. Information thus travels from sensory receptors → thalamus → primary cortex → secondary cortex → tertiary cortex (Kolb and Whishaw 2009).
2. Serial processing is hierarchical, with increased complexity at each level.
3. "Our perceptions of the world are unified and coherent entities" (Kolb and Whishaw 2009, p. 267).

Stated by Luria, these three units represent (1) muscle tone/walking, (2) processing/storing external information, and (3) programming/regulating/verifying mental activity (Luria 1973a, p. 43). Luria argued that every form of "conscious" activity was an output of the combined efforts of these three functional units working in unison (Luria 1973a). Luria was very fascinated by studying the interactions and specific contributions of these three functional units (Luria 1973a).

Subsequently, Daniel Felleman and David Van Essen would further refine this hierarchical model to include hierarchically organized cortical areas but with more than "one area occupying specific position relative to other areas, but with more than one area allowed to occupy a given hierarchical level" (Kolb and Whishaw 2009, p. 267). They coined this model a "distributed hierarchical system" in which the number of areas expands as we go up the hierarchy (Kolb and Whishaw 2009).

Beyond functional units, Luria provided an explanation for subdivisions he called zones (Luria 1973a). These zones are known as the primary, secondary, and tertiary zones. The primary zone is primarily responsible for units utilizing neurons to receive impulses from sensory organs. Put differently, the primary zone in the motor strip of the frontal lobe focuses on motor output (Kolb and Whishaw 2009; S. Goldstein, personal communication, July 13, 2013). The second zone has neurons employed to enable incoming excitation to be moved along to the tertiary zones (Luria 1973a). This zone sequences motor activity and speech production, for example. This tertiary zone is then responsible for the organization and integration of excitation received by sensory structures and rearranging stimuli into a linear order (Luria 1973a). Damage to the frontal regions is particularly troublesome in altering the behaviors and executive functions of individuals, due to its influence and connections with other cortical and subcortical areas of the brain including the thalamic, hypothalamic, and limbic systems (Kolb and Whishaw 2009).

Luria would focus a large portion of his career to understand the contribution made by units and zones of the brain to understand which mental activities are encompassed under the umbrella of which brain systems (Luria 1973b, p. 103). This is highlighted by one of Luria's salient quotes: "Finally I discussed the chief sources of our knowledge of the cerebral basis of mental activity and I showed that of these three sources—the comparative anatomy of the brain, methods of stimulation and methods of destruction of its individual areas—with respect to the analysis of the functional organization of the human brain it is the last which is evidently the most important" (Luria 1973b, p. 103).

Disturbance of Higher Mental Functions

Thus far, we have focused much of our effort toward describing the "function" from Luria's perspective. It is appropriate to now turn to dysfunction from Luria's perspective. As Luria has stated, "In order to learn more about human cognitive functions we must study both their unfolding and disruption" (Kaczmarek 1999).

One cannot assume that studying brain lesions will provide comprehensive, generalizable answers. Inherently, brain lesions destroy various areas of the brain, “not just one narrowly localized group of nerve cells” (Luria 1973b, p. 104). “The initial hopes for using disturbance of higher mental functions for local diagnosis of brain damage began to appear unfounded, and the possibility of using psychological symptoms for local diagnosis became a controversial issue” (Luria 1964, p. 4). Expounded by Luria, the brain is now functioning under “pathologically changed conditions,” meaning that some elements of the brain may be completely destroyed while others function as normal—making it very difficult to study the brain of those under “pathologically changed conditions” and apply this to the neurotypical population or even to other individuals with brain injuries (Luria 1973b, p. 104). Furthermore, “Nothing can be more mistaken than such an idea and such an attempt to localize the symptom of apraxia (and consequently the function of praxis) in a narrow area of the cortex” (Luria 1973b, p. 35). And, “As we have seen, a local brain lesion does not lead to the direct loss of a particular mental condition; this was the view held by the supporters of narrow localizations” (Luria 1973b, p. 103).

Luria believed that any sustained damage to a pinpointed cortical area in earlier childhood would result in “a relatively elementary basis of mental activity, and it unavoidably causes a secondary ‘systemic’ effect, or an underdevelopment of higher structures built on these elementary functions” (Luria 1973b; Kotik-Friedgut 2006, p. 48).

However, in adults, Luria argued, damage to “higher zones” produces just the opposite effect—elementary functions would deteriorate depending on the secondary or higher forms of activities (Luria 1973a; Kotik-Friedgut 2006, p. 48). However, he did not suggest that higher mental functions are “built up as a second story over elementary processes. Rather, higher mental functions are the product of the marriage of elementary functions into the new system under new guidance” (Luria 1973b, p. 43).

One distinction between humans and animals applies to the frontal lobes of the brain—the last features of the cerebral hemispheres to be completely formed. Primates possess larger frontal lobes than other animals, maturing between 4 and 7 years of age in humans (Luria 1973a, p. 187). The distinction between the frontal lobes of primates and non-primates lies within voluntary attention. After all, “Man lives in a constantly changing environment, and these changes, which are sometimes unexpected by the individual, require a certain level of increased alertness” (Luria 1973a, p. 55). A disruption to the frontal lobes may lead to “substantial disturbances in the flow of intellectual processes” (Luria and Tsvetkova 1964, p. 97). “These patients do not make a programme, there is no consequent realization of an original plan, and actions take on the character of change trials, which easily fall under the influence of immediate impressions or perseverations. If the results achieved do not match with the original intention, the mistakes made are not recognized and not corrected” (Luria and Tsvetkova 1964, p. 107).

Along with aforementioned case studies derived from the aftermath of World War II, Luria and colleagues investigated a patient with a verified arachnoidal endothelioma (meningioma) at the Burdenko Institute of Neuropsychology, patient Zav in 1962 (Luria et al. 1964). With complaints of forgetfulness, headaches, and nausea, Zav had difficulties following through with instructions during her neuropsychological evaluation. Once language was more complex, her speech behavior was replaced by perseveration that disrupted her when attempting to answer questions (Luria et al. 1964). Luria and colleagues provided a wealth of information regarding Zav’s progression through extensive neuropsychological examination. “Such deviation from the assigned instructions and the substitution of a simple response was shown with simultaneous complete retention of the instruction in the patient’s verbal system. As with the other tasks, the process of relating her actions to the instruction was hampered. Recognition of errors was almost impossible” (Luria et al. 1964, p. 267).

Overall, much of Zav’s ability to carry out actions was dependent upon the complexity of the instructions (Luria et al. 1964). “At first glance one may suppose that they preserve all the basic functions of the Human Brain. But this is not the case. Attentive observation shows how deep are the disturbances in the regulation and control of the conscious behavior of these patients" (Luria 1969b, p. 12). Further, "The suggestion, which has been detailed elsewhere, is that frontal lobe resection limits the subject's ability to organize and reorganize his behavior when flexibility is demanded, and especially when external cues to support such organization are wanting …the patient with frontal lobe destruction is often unable to fulfill instructions, is unable to inhibit impulsive reactions or to hold back the tendency towards fixed repetition of movement" (Luria et al. 1964, p. 258).

Luria postulated, "This means that symptoms of disturbance of any higher cortical function may be used for local diagnosis of brain damage, but that such a diagnosis can be carried out only when there is qualitative analysis or evaluation of the symptom. Such evaluation of the symptom is the fundamental task of neuropsychology" (Luria 1964, p. 13).

A Synopsis of Luria in the Final Decade

Luria corresponded with colleagues throughout the world, questioning their conceptualization and understanding of the extra-cortical organization of brain functions (Kotik-Friedgut 2006). Six years before his death, Luria corresponded with Professor Douglas Bowden in the United States. This correspondence, dated February 20, 1971, from Souhanovo, Russia, was included in Dr. Bowden's *Meta-principles in Luria's Neuropsychology*. Excerpts of this correspondence appear below (adopted and transcribed as it appears in Dr. Bowden's work). These excerpts very well summarize Luria's thoughts in his final years:

1. There are higher cortical (or psychological) functions specific to human being and not existing in animals. These specifically human psychological processes (or functions) derive from social sources, i.e., tool-using social behavior of man.
2. The most specific feature of these higher processes is that they are tool- or means-using processes. Animals do not use means, their behavior is not mediated by means, tools, or signs, and it is all unmediated, natural behavior, whereas human behavior is always mediated (tool- or means-using) or indirect structure, social by origin and voluntary (or conscious) by the modes of work.
3. This indirect, means-using behavior is mediated via speech—the most important system of tools or signs in human history. Language (or speech) has not only its semantic function, function of categorization of impressions, but its pragmatic or regulatory (or controlling) functions as well. By using language, man overcomes the direct influences of environment, and his behavior becomes no more field linked but is goal or plan linked.
4. The indirect, tool (sign)-using behavior starts a new form of cortical work: human cortex is no more a complex of work of different zones, organized by influences of the centrencephalic system (or natural drives): it becomes a historically organized, plastic functional system where language plays a decisive organizing role. That is why higher cortical functions of man have to be evaluated as functional systems of cortical zones, linked by the leading role of language as a decisive means of behavior.
5. The disorganization of the brain functions following local brain lesions is in no way a partial deficit (destruction of a spherical local function) and no more a total lowering of the general brain activity. It results in a disorganization of functional brain systems, each time resulting from a defect of a *basic factor* (according the locus of the lesion)—bringing series of primary symptoms and resulting in a series of *secondary symptoms* or functional (systemic) results.
6. The basic goal of neuropsychology is neither a pure *description* nor a direct reduction to a physiological issue but careful *psychological qualification* of the symptom (i.e., singling out the underlying factor and then a description of systemic results of the destruction or elimination of this factor), that is, the real way to the neuropsychological diagnostics of brain injury.

A Post-Luria World

After laboring in the field of neuropsychology for much of his adult life, Luria passed away on August 14, 1977, at the age of 75 from cardiac arrest. Luria's research continued directly and indirectly through his pupil named Alfred Ardila among others (Cole et al. 2006). Continuing work under the tenets of cross-cultural neuropsychology, Ardila is one of many scholars perpetuating the work of Alexander Luria (Kotik-Friedgut 2006). A centennial celebration of Luria's birth was not necessary to broadcast his true impact on the world. His contributions to the field of neuropsychology are immeasurable. Our current conceptualization of extra-cortical organization of higher mental functions is one of Luria's most notable principles (Kotik-Friedgut 2006).

Because of Luria's influence, the field of neuropsychology has progressed rapidly in recent years. He shaped the process of learning how we think, learn, and solve problems. "In terms of assessment techniques, Luria's methods are qualitative and flexible; he seeks links in functional systems, his methods are clinical-theoretical and case oriented" (Cole 2005a, p. 35). By contrast, North American neuropsychologists rely on psychometric, actuarial, quantitative, group studies" (Cole 2005a, p. 35). Furthermore, "By the early 1980s, neuropsychology was no longer confined to a few elite laboratories, and the new field of clinical neuropsychology blossomed in the clinics and hospitals. Since that time, three factors have enhanced the rate of change in neuropsychological assessment: functional imaging, cognitive neuroscience, and managed healthcare" (Kolb and Whishaw 2009, p. 806).

One of the many assessment tools in the field of neuropsychology today founded on Luria's theories is the Cognitive Assessment System (Naglieri & Das, 1997). The CAS was developed by J.P. Das and Jack Naglieri. The CAS attempts to measure the functional units described by Luria. Employed with children aged 5–17, the CAS tracks children as they mature, measuring their executive functioning through the years (Das 1980; S. Goldstein, personal communication, July 13, 2013). The functional dimensions of brain structures as described by Luria are described in terms of planning, attention, and simultaneous and successive processes or abilities (Luria 1966, 1973b, 1980; Das and Naglieri 1997). More popularly known as the PASS theory, these processes are based on Luria's neurodevelopmental model of stages of higher maturation (S. Goldstein, personal communication, July 13, 2013). The CAS has recently been revised for a second edition.

We previously outlined the three functional units, which apply to PASS theory: attention uses the first unit (i.e., midbrain, medulla, thalamus, and hypothalamus), simultaneous and successive the second, and planning the third (Luria 1973b; Koxiol 2009; S. Goldstein, personal communication, July 13, 2013). Simultaneous processing refers to the integration of stimuli into interrelated groups or a whole (e.g., following multistep instructions) (Das 1994). Successive processing involves sequencing of stimuli in a serial order (e.g., decoding unfamiliar words and speech articulation) (Das 1994). In a sense, the roles of successive and simultaneous processing are reversed. Planning, associated with the first unit, is a process in which an individual evaluates solutions to problems. Attention, also associated with the first function, entails a process in which an individual focuses on particular stimuli (Das 1994).

Concluding Remarks

Alexander Luria was active in the field of neuropsychology up to his passing. His work would not be forgotten. His unfinished publications would later be published by colleagues as *Paradoxes of Memory* (Cole et al. 2006). "Depicted by four approaches, Luria outlined the derivation of human characteristics, human consciousness, psychology as a biological science, and psychology as a unique science" (Cole 2005a, p. 37). In perhaps his last published article before his death, Luria continued his craft in both identifying and distinguishing psychological approaches to better confirm the tenets of his own philosophies (Cole 2005a, p. 35). One of the most poignant statements made by Luria on this subject was, "As I have said already, any human mental activity is a complex functional system effected through a combination of concertedly working brain structures, each of which makes its own contribution to the functional system as a whole" (Luria 1973b, p. 38).

In the famous words of Luria before his death in 1977: "In this respect neuropsychology is merely the most complex and newest chapter of neurology, and without this chapter, modern clinical neurology will be unable to exist and develop. Neuropsychology now has a firm foothold in clinical neurology and neurosurgery. This fact is a source of great satisfaction to the author who, together with his colleagues, has spent a good deal of his life in an effort to make neuropsychology an important practical branch of neurology. It gives him confidence that his scientific life has not been spent in vain, and that new and important prospects lie ahead for neurology, and for those important divisions of neurology—the topical diagnosis of local brain lesions and the rehabilitation of these patients" (Luria 1973a, p. 344).

Those familiar with Luria's work would likely express immense gratitude for his contributions and assure him that his scientific life was surely not spent in vain. Close disciples of Luria's work cannot begin to measure the salience of Luria's work as it relates to the current times.

The modern tenets of neuropsychology are built upon Luria's teachings, much as a religion's philosophies are built upon a book of God, such as the Bible. As Luria discussed the ever-changing climate of an individual's environment, so too, the climate of neuropsychology will continue to evolve. Having said that, the basic tenets of neuropsychology as provided by Luria will prevail in secure establishment as *the* foundation of neuropsychology. As stated by Luria a few years prior to his death, "Thus we are still very far from the solution of our basic problem—the Neuropsychological Organization of Man's Conscious Action, and we can only look forward with envy and hope to the work of the next generations of Psychologists who will one day take our place and bring to a successful end the work we have only started" (Luria 1969a, p. 20).

References

Akhutina T.V., & Pylaeva N.M. (2011). L.Vygotsky, A. Luria and Developmental Neuropsychology. Psychology in Russia: State of the Art(4), 155–175.

Bowden, D. M. (1971). Keystone to Luria's neuropsychology. *Skolepsykologi, 8*(6), 409–416.

Cole, M. (2005). A. R. Luria and the cultural-historical approach in psychology. In T. Akhutina, J. Glozman, L. Moskovich, & D. Robbins (Eds.), *A. R. Luria and contemporary psychology: Festschrift celebrating the centennial of the birth of A. R. Luria* (pp. 35–41). New York: Nova Science Publishers.

Cole (2005b) Alexander Luria. Retrieved from http://luria.ucsd.edu/index.html

Cole, M., Levitin, K., & Luria, A. (2006). *The autobiography of Alexander Luria: A dialogue with the making of mind*. Mahwah: Lawrence Erlbaum Associates.

Das, J. P. (1980). Planning: Theoretical considerations and empirical evidence. *Psychological Research (W. Germany), 41*, 141–151.

Das, J. P. (1994). Serial and parallel processing. Encyclopedia of Human Intelligence, 964–966.

Kaczmarek, B. L. J. (1999). Extension of Luria's psycholinguistic studies in Poland. *Neuropsychology Review, 9*(2), 79–87.

Kolb, B., & Whishaw, I. Q. (2009). *Fundamentals of neuropsychology*. New York: Worth Publishers.

Kotik-Friedgut, B. (2006). Development of the Lurian approach: A cultural neurolinguistic perspective. *Neuropsychology Review, 16*, 43–52.

Koxiol, L.F., & Budding D.E. (2009). *Subcortical structures and cognition: Implications for neuropsychological assessment*. New York: Springer Books.

Luria, A. (1932a). *The nature of human conflicts*. Reprint. New York: Liverlight, 1976.

Luria, A. R. (1932b). *The nature of human conflicts—or emotion, conflict, and will: An objective study of disorganisation and control of human behaviour*. New York: Liveright Publishers.

Luria, A. R. (1959). Disorders of 'simultaneous perception' in a case of bilateral occipito-parietal brain injury. *Brain, 82*(3), 437–449.

Luria, A. R. (1963). *Restoration of function after brain injury*. Oxford: Pergamon Press.

Luria, A. R. (1964). Factors and forms of aphasia. Reprinted from Ciba Foundation Symposium on Disorders of Language, 1964, pp. 143–161.

Luria, A. R. (1968). The mind of a mnemonist (trans. Solotaroff, L.). New York: Basic Books.

Luria, A.R. (1969a). The origin and cerebral organization of man's conscious action: an evening lecture to the XIX International Congress of Psychology. London.

Luria, A. R. (1969b). Neuropsychology in the local diagnosis of brain damage. Reprinted from Cortex, Volume 1, 1964, pp. 3–18.

Luria, A. R. (1970). *Traumatic aphasia: Its syndromes, psychology, and treatment*. The Hague: Mouton de Gruyter.

Luria, A.R. (1972). *The man with a shattered world (L. Solotaroff, Trans.)*. New York: Basic Books.

Luria, A. R. (1973a). *The working brain: An introduction to neuropsychology* (trans: Haigh, Basil). London: Penguin Books Ltd.

Luria, A. R. (1973b). *The working brain*. New York: Basic Books.

Luria, A. R. (1976). *The cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books. (Original work published 1966.)

Luria, A. R., Cole, M., & Cole, S. (1979). *The making of mind: A personal account of Soviet psychology*. Cambridge, Mass: Harvard University Press.

Luria, A. R. (1987). The man with a shattered world: The history of a brain wound. Cambridge, Mass: Harvard University Press.

Luria, A. R., & Tsvetkova, L. S. (1964). The programming of constructive activity in local brain injuries. *Neuropsychologia, 2*, 95–107.

Luria, A. R., Pribram, K. H., & Homskaya, E. D. (1964). An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). *Neuropsychologia, 2*, 257–280.

McCloskey, G., Perkins, L., Van Divner, B. (2009). *Assessment and Intervention for Executive Function Difficulties*. Routledge, New York.

Naglieri, J. A., & Das, J. P. (1997). Cognitive Assessment System. Administration and scoring manual. Interpretive handbook. Itasca, IL: Riverside.

Semrud-Clikeman, M. & Fine, J. (2008). *Brief and comprehensive neuropsychological assessment of children and adults with learning difficulties and disabilities*. In R.D'Amato (Ed.). Neuropsychology of learning (pp. 245–291). Mahway, NJ: Lawrence Erlbaum.

Vygotsky, L. S., Veer, R., & Valsiner, J. (1994). *The Vygotsky reader*. Oxford, UK: Blackwell.

Intelligence: Defined as Neurocognitive Processing

14

Tulio M. Otero

Introduction

Intelligence does not exist. Although this is a strange statement, it is a fundamental truth that must be understood before understanding what intelligence is and how we define it in this chapter. Intelligence is not a tangible entity that can be measured in the same way physicists measure matter. Rather, intelligence is a hypothesized "phenomenon." Its ontology, etiology, and scale are inferred through indirect means. We assume that intelligence is found in the brain, yet we cannot locate it. It is not a discrete, embodied force radiating from the brain that can be measured directly with sophisticated apparatus. There is nothing physical to set a ruler next to and claim, "this is how much intelligence is here" (e.g., Thorndike 1997).

Human intelligence is a psychological construct defined by many influential philosophers, researchers, and theoreticians at different points in history. Thorndike (1997) described a construct as a defensible collection of separate, quantifiable qualities and attributes that, when taken together, form a measurable exemplification of a multifaceted, hypothesized abstraction. Considering the definitions of intelligence as a construct is of significant importance. Clinicians and researchers may define intelligence in dissimilar ways. As a result, their findings may produce conflicting results. Failure to consider the underlying definition of intelligence makes accurate interpretation of collected data a tenuous process. Pluker and Esping (2014) present definitions of intelligence by 12 prominent historical and 7 contemporary theorists. Of interest is that only one researcher and theoretician, JP Das, defined intelligence as the "sum of all cognitive processes" (p. 19).

Intelligence as used in this chapter refers to a subset of psychological processes that involve cognition. Neuropsychological and neuroanatomical studies have elucidated that neuropsychological assets and deficits in cognitive processes can facilitate or impair specific types of learning and performance. An examination of anatomy also illustrates how and why some neurocognitive processes are interrelated. Basic psychological processes are important for any individual to be able to interact effectively with the environment, to learn from formal instruction and from experience, and to adapt to new situations. Our brain takes in and processes new information in such a manner that typically the right frontal lobe systems are engaged until new information is learned, assimilated, and accommodated into our existing knowledge base or becomes routinized (Goldberg 2009). Once information or knowledge becomes readily familiar, greater activation of the left frontal systems appears to come "online" for handing this

T.M. Otero, Ph.D. (✉)
The Chicago School of Professional Psychology,
325 N. Wells St., Chicago, IL 60654, USA
e-mail: TOtero@thechicagoschool.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_14,

and is known as the novelty-routinization principle (Goldberg 2009). This novelty-routinization principle exemplifies a dynamic view of the brain function that provides us with one way of understanding the changing underlying neuro-anatomical and neuropsychological events that occur during learning and skill development (Koziol 2014). These hemispheric asymmetries are thought to support different styles of information processing.

The right hemisphere processes information in a simultaneous holistic manner, whereas the left hemisphere processes information in an analytic, successive manner. This allows for the acquisition of harmonizing information about the world. Hence, these two modes of processing provide more robust information, different than what would be obtained from one type of processing in isolation. The brain is far more complex than this simple dichotomy, however. The great neuropsychologist Alexander Luria, in several of his writings (Luria 1966, 1980, 1982), maintained that the brain is complex and that no part of it functions without the cooperation of other parts. Thus, Luria viewed the brain as a functional mosaic, meaning that various parts interact in different combinations to apply varying combinations of cognitive processing abilities (Luria 1973). Thus, Luria contended that there is no area of the brain that functions without input from other areas. Integration of processing abilities is a key principle of brain function within the Lurian framework.

During the past decade, ideas about the functional specialization of brain regions have dramatically evolved. According to Johnson and colleagues (2005) and Friston (2002), functional specialization can be defined as the degree of information processing specificity of an identified brain region for a particular cognitive ability or facet of cognitive operations. However, as Luria pointed out, brain regions obviously do not function in isolation. The functional architecture of the brain is characterized by reciprocal connectional brain profiles of the cerebro-cortical, cortical-basal ganglia, cerebro-cerebellar, and basal ganglia-cerebellar circuitry systems (Bostan, A. C., Dum, R. P., & Strick, P. L. 2010, 2013; Bostan and Strick 2010; Koziol et al. 2011).

The importance of understanding neurocognitive processing allows us to understand not only the dynamic nature of the brain but also allows us to understand differences in learning and skill development. Technologies such as magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), computerized tomography (CT), and diffusion tensor imaging (DTI) have reduced the need for neuropsychological tests to localize and access brain damage. Although these technologies are favored for investigating the structural and functional dynamics of the brain, it is these authors' opinion that the understanding and assessment of neurocognitive processes by studying patterns of neurocognitive strengths and weaknesses in developmental, psychiatric, psychosocial, and learning disorders are best achieved through formal assessment procedures. By addressing both brain functions and environmental factors intrinsic in complex behaviors, such as thinking, reasoning, planning, and the variety of executive capacities, clinicians are able to offer needed services to children with a variety of learning, psychiatric, and developmental disorders. Several neuropsychological tests play an important role in identifying the neurocognitive processes, or abilities, necessary for effective thinking, learning, and behaving, while also allowing for judgments regarding the integrity of the brain. Psychologists of different specialties may use standardized instruments to collect information and derive inferences about brain-behavior relationships. Traditional neurocognitive testing and evaluation takes a cortical-centric approach to understanding brain-behavior relationships. Neuropsychological tests can also be utilized as one way to assess the integrity of cortical-subcortical functional networks such as the fronto-striatal system, among others (Koziol 2009). Consistent with current functional conceptualizations of the brain, we believe both cortical and subcortical networks are important for basic neuropsychological processes to manifest efficiently.

A conceptualization of human cognitive functioning like the one described by A. R. Luria can guide the development of assessment tools. Such tools should not only evaluate the underlying neurocognitive processes necessary for efficient thinking and behavior but also provide for the development of effective interventions and address the question of prognosis.

Neuropsychological Theory and PASS Processes

Luria's theoretical account of dynamic brain function is perhaps one of the most complete (Lewandowski and Scott 2008) and current research in the area of large neuro-networks continue to lend support to his original observations (Koziol 2014). Luria conceptualized four unified levels of brain-behavior relationships and neurocognitive disorders that the clinician needs to know: the structure of the brain, the functional organization based on structure, syndromes and impairments arising in brain disorders, and clinical methods of assessment (Korkman 1999). His theoretical formulations, methods, and ideas are articulated in works such as *Higher Cortical Functions of Man* (1966, 1980) and *The Working Brain* (1973). Luria viewed the brain as a functional medley, the parts of which interact in different combinations to subserve cognitive processing (Luria 1973). Cognition and behavior then result from an interaction of complex brain activity across various areas. Luria's (1966, 1973, 1980) research on the functional aspects of brain structures formed the basis for the development of the PASS theory (planning, attention, simultaneous, successive processing), initially described by Das et al. (1994) and operationalized by Naglieri and Das (1997a, b) in the *Cognitive Assessment System* (CAS), and most recently in the Cognitive Assessment System-Second Edition (Naglieri et al. 2014a, b).

From a Lurian framework, cognitive functions, such as attention, executive functions, language, sensory perception, motor function, visuospatial facilities, and learning and memory, are multifaceted capacities. They are composed of flexible and interactive subcomponents that are mediated by equally flexible, interactive, neural networks (Luria 1962, 1980). These cognitive functions are theorized as three separate but connected "functional units" that provide four basic psychological processes. The three brain systems are referred to as "functional" units because the neuropsychological mechanisms work in separate but interrelated systems. In other words, multiple brain systems mediate complex cognitive functions. For example, multiple brain regions interact to mediate attentional processes (Koziol, Joyce, & Wurglitz, 2014). The executive functions subserved by the third functional unit, as described by Luria, regulate the attentional processes of the first functional unit in sustaining the appropriate level of arousal and vigilance necessary for the detection and selection of relevant details from the environment. Consider the case of response inhibition; the executive function of inhibition allows a student to resist or inhibit responding to salient, but irrelevant, details on a task. Response inhibition allows the student to sustain focus, over time, on task relevant features.

The brain systems described above are consistent with the four psychological processes identified by the PASS theory (Naglieri and Das 1997b), and this amalgamation of processing abilities is a key principle of brain function within the Lurian framework. Cognition and behavior result from an interaction of complex brain activity across various areas. Naglieri and Das (1997a, b) used Luria's work as a base to redefine intelligence from a multi-ability perspective. The PASS theory has strong empirical support (see Das et al. 1979, 1994b) and since the publication of the Cognitive Assessment System (CAS) and Cognitive Assessment System-Second Edition (CAS-2) (see Naglieri 2012; Naglieri and Conway 2009; Naglieri and Otero 2011).

Luria (1973) stated "each form of conscious activity is always a complex functional system and takes place through the combined working of all three brain units, each of which makes its own contribution" (p. 99). In other words, the four processes form a "working constellation" (Luria 1966, p. 70) of cognitive activity. Thus, a child or adult can use different combinations of the four psychological processes in conjunction with their knowledge and skills to perform a task. Although effective functioning is achieved through the appropriate combination of all processes as demanded by the task, each process is not equally involved in every task. For example, reading comprehension may predominately involve one process, while reading decoding can be strongly dominated by another (Das et al. 1994b). Or basic math calculation may require more of one process, while math-reasoning tasks may require a different cognitive process. For example, learning basic math calculation operations may initially require a basic step-by-step approach. Understanding math-reasoning problems, however, requires holding multiple elements of the task in memory, surveying the elements, and making decisions about these elements before solving the problem. Das and Naglieri and their colleagues used Luria's work as a blueprint for defining the basic neuropsychological processes that underlie human performance (Naglieri 2003). Their efforts represent the first time that a specific researched neuropsychological theory was used to provide an alternative conceptualization of human intelligence.

Three Functional Units Described

Luria (1973) provided considerable evidence for the neuropsychological processes associated with each of the three functional units and their association with specific regions of the brain. Briefly stated, these three functional units have been used by Naglieri and Das (1997b) as the basis of planning (third functional unit), attention (first unit), and simultaneous and successive (second unit) cognitive processes. The brain stem, the diencephalon, and the medial regions of the cortex are the primary locations for the first of the three functional units of the brain, the attention-arousal system, (Luria 1973). This unit is comprised specifically of the midbrain, medulla, thalamus, and hypothalamus. These structures work in concert to maintain the appropriate cortical tone. Recent formulations of these regions suggest some structures at the level of diencephalic and medial regions have reciprocal connections to the cortex through a variety of large-scale brain circuitries (Koziol and Stevens 2012) potentially influencing a wide range of behaviors (Koziol 2009).

First Functional Unit

Attention is a basic component of intelligent behavior involving allocation of resources and effort. Arousal, attention, effort, and capacity are concepts that have a complex relationship and importance for understanding behavior. When a person is required to pay attention to only one dimension of a multidimensional stimulus array, the inhibition of responding to other (often more salient) stimuli and the allocation of attention to the central dimension are required. Luria stated that optimal conditions of arousal are needed before the more complex forms of attention involving "selective recognition of a particular stimulus and inhibition of responses to irrelevant stimuli" (Luria 1973, p. 271) can occur. This way of conceptualizing attention is analogous to such contemporary models as Mirsky and Duncan's (2001, 2003; Koziol 2014) in which focus, shift, sustenance, and stabilization of attention are necessary before complex learning can take place. Moreover, the second and third functional units can operate effectively only after individuals are sufficiently aroused and their attention is adequately focused.

Second Functional Unit

The occipital, parietal, and temporal (particularly medial temporal portions) lobes posterior to the central sulcus of the brain are associated with the second functional unit. Information from the external world is received, processed, and retained within this unit. Thus, the major function of the second functional unit is sensory reception and integration (Semrud-Clikeman and Teeter Ellison 2009). The areas of the second functional unit correspond to their sensory modality: temporal for auditory stimuli, parietal for tactile, and occipital for visual. This functional unit is hypothesized to be guided by three functional laws: (1) during development, the makeup of cortical zones does not remain the same; (2) the specificity of function of the cortical zones decreases with development; and (3) an increase in lateralization of function increases with development (Luria 1980). This hierarchy is further subdivided into *zones*. The primary zones of these structural units employ modality-specific groups of neurons to receive impulses from the sensory organs, while the secondary zones of these structures surround the primary zones with associative neurons which enable incoming excitation to be conveyed to the tertiary zones. These tertiary zones are responsible for integrating and organizing the excitation arriving from the different sensory structures and converting the stimuli which are received in a specific linear order into simultaneously processed groups (Luria 1974). The secondary zones are involved in the input of data and integration of information. These zones process information sequentially and connect cross-modally with several stimuli impinging on the brain at a time. For example, reading is an integration of both visual and auditory material, and mathematics is the integration of visual material with the knowledge of numbers and quantity.

There are several secondary zones for auditory, tactile, and visual information. The auditory secondary zone lies within the secondary regions within the temporal lobes and involves the analysis and synthesis of sounds and the sequential analysis of phonemes, pitch, tone, and rhythm. The secondary tactile zone is within the parietal lobe and is involved in the recognition of complex tactile stimuli and two-point discrimination, for example. The secondary visual zone borders the primary visual cortex of the occipital lobe. Visual discrimination of letters, shapes, and figures are related to it. Traditional intelligence tests are hypothesized to measure some aspects of the second functional unit (Semrud-Clikeman and Teeter Ellison 2009). Because the second functional unit can be considered as the center for analysis, coding, and storage of information, damage to the structures forming the second functional unit can result in difficulty across all academic areas.

Tertiary zones process and integrate information from all sensory areas. This integration of information from various modalities occurs through simultaneous processing. For example, some math involves the integration of both visual materials and knowledge of number quantity. Math reasoning, in the form of word problems, may additionally involve the integration of grammatical skills, analysis of auditory information, and the comprehension of auditory or written material. Damage to these zones has been related to lower measured intelligence and difficulties across several basic academic areas.

Simultaneous and successive processing are subserved by the second functional unit. Simultaneous processing is a mental activity by which a person integrates stimuli into interrelated groups or a whole. For example, in order to follow multistep directions, the relationships among the different parts of what is said must be correctly understood. Reading unfamiliar words that are initially difficult to decode and then are later quickly and effortlessly recognized as a whole word is another example of a task initially demanding successive processing, followed by the efficient simultaneous processes of reading the word as a whole unit. Children presenting with difficulties performing on tasks that require learning new information by associating it with other information may have deficits in simultaneous processing. Difficulty integrating visual information may be a primary deficit in children with nonverbal learning disabilities. These children have been found to do more poorly on measures that demand simultaneous processing such as tests on visual motor integration and visual-perceptual skills compared to children with ADHD and normally developing children (Wilkinson and Sermund-Clikeman 2008).

While simultaneous processing involves working with interrelated stimuli, successive processing *requires* work with stimuli in a specific serial order. This processing ability is required when a child arranges things in a strictly defined order, where each element is only related to those that precede it and these stimuli are not interrelated.

Successive processing ability involves both the perception of stimuli in sequence and the formation of sounds and movements in order. For example, successive processing is involved in the decoding of unfamiliar words, production of syntactic aspects of language, and speech articulation. Following a sequence such as the order of operations in a math problem is another example of successive processing. Most real-life situations will require both of these processes to become activated to some degree.

Third Functional Unit

One of the most extraordinary capacities we have as humans is our ability to reflect and self-direct behavior. This ability is frequently described using the term executive function (EF). Although many definitions of EF abound (Goldstein et al. 2014), the concept of EF is intimately linked to the frontal lobes. The prefrontal areas of the frontal lobes of the brain are associated with the third functional unit (Luria 1980). The prefrontal cortex is well connected with every distinct functional unit of the brain (Goldberg 2009). This unit is mostly responsible for output planning and with most behaviors we typically consider as executive functions. The third functional unit is also further differentiated into three zones, with the primary zone in the motor strip of the frontal lobe being concerned with motor output. The secondary zone is responsible for the sequencing of motor activity and speech production, while the tertiary zone is primarily involved with behaviors typically described as executive functions. Damage to any of several areas of the frontal regions has been related to difficulties with impulse control, learning from one's mistakes, delay of gratification, and attention. Because the third functional unit has rich connections with other parts of the brain, both cortical and subcortical, there can be forward and backward influences, to and from other regions such as the cerebellar, thalamic, hypothalamic, and limbic areas. Additionally, a growing body of evidence points to a network of connected regions in the adjacent frontal and parietal lobes, which have been implicated in higher-order processing such as attention, decision-making, and intelligence (Kolb and Whishaw 2009).

Luria stated that "the frontal lobes synthesize the information about the outside world … and are the means whereby the behavior of the organism is regulated in conformity with the effect produced by its actions" (Luria 1980, p. 263). The frontal lobes provide for the programming, regulation, and evaluation of behavior and enable a person to ask questions, develop strategies, and self-monitor (Luria 1973). Other responsibilities of the third functional unit include the regulation of voluntary activity, conscious impulse control, and various linguistic skills such as spontaneous conversation. The third functional unit provides for the most multifaceted aspects of human behavior, including personality and consciousness (Das 1980). The first and third functional units share a reciprocal relationship. The higher cortical systems both regulate and work in collaboration with the first functional unit while also receiving and processing information from the external world and determining an individual's dynamic activity (Luria 1973). It is both influenced by the regulatory effects of the cortex and influences the tone of the cortex. The ascending and descending systems of the reticular formation enable this relationship by transmitting impulses from the lower parts of the brain to the cortex and vice versa (Luria 1973). Thus, damage to the prefrontal area can alter this reciprocal relationship, so that the brain may not be sufficiently aroused for complex behaviors requiring sustained attention. A breakdown in any portion of the complex loop-like interactions between the prefrontal, ventral brain stem and posterior cortex is likely to produce symptoms of attention deficit (Goldberg 2009).

The psychological processes that are routed in each of the functional units are linked. For the PASS theory, this means that psychological processes of attention and planning are strongly related because planning often has conscious control of attention. In other words, one's limited attentional resources are dictated by the plan for one's behavior. However, attention as well as the other PASS processes are influenced by many variables other than planning. One of these influences is the environment. Novel encounters within daily life demand we act in one way or another. Several PASS processes can be involved as we make judgments about similarities and differences between past situations and the present demands, while hypothesizing possible outcomes of our actions and as we select behaviors while acting on the environment.

Functional Units: Interactions and Influences

Luria believed no part of the brain works by itself, and therefore, his organization of the brain into functional units was not an attempt to map out the precise locations where specific areas of higher cognition took place, but rather to emphasize that no cognitive task solely requires simultaneous, successive, planning, or attention processing or any other process, but rather, it is a matter of emphasis. He stated, "…perception and memorizing, gnosis and praxis, speech and thinking, writing, reading and arithmetic, cannot be regarded as isolated or even indivisible 'faculties'…" (Luria 1973, p. 29). That is, an attempt to identify a fixed cortical locale for any complex behavior is a mistaken endeavor. Instead, the brain should be conceptualized as a functioning whole comprised of units that provide purpose.

Activities such as reading and writing can be evaluated and seen as constellations of activities related to specific working zones of the brain that support them (Luria 1979). This means that since the brain operates as an integrated functional system, even a minor disruption in an area can cause disorganization in the entire functional system (Varnhagen and Das 1986). Thus, many behaviors may be impacted by a disruption caused by a lesion, damage, or underdeveloped structures. For example, lesions or damage to the prefrontal cortex, with its complex connections with other areas of the brain including several subcortical areas, may result in affective dissociations, impaired executive functions, poor judgment and processing, or intellectual deficits.

Luria believed that a child's cultural experience is a significant influence on the functional units and also a necessary foundation that aids the development of human cognition (Luria 1979). The organization of the brain into functional units also accounts for the interaction of cultural influences and biological factors within higher cognition. Luria (1979) notes "…the child learns to organize his memory and to bring it under voluntary control through the use of the mental tools of his culture" (p. 83). Kolb et al. (2003) also wrote that although "the brain was once seen as a rather static organ, it is now clear that the organization of brain circuitry is constantly changing as a function of experience" (p. 1). Various brain systems are highly modifiable by experience and dependent on experience only during particularly sensitive time periods and other systems remain capable of change by experience throughout life (Neville 2006; Neville and Stevens 2008). Similarly, Vygotsky (1976) described this interplay when he described speech as a self-regulatory function. Self-talk functions as self-guidance and regulation, helps children think about their mental activities and behaviors, and select courses of action, and is the foundation for all higher cognitive processes (e.g., controlled attention, deliberate memorization and recall, categorization, planning, problem-solving, abstract reasoning, self-reflection). Stuss and Benson (1990) described this interplay as follows:

> The adult regulates the child's behavior by command, inhibiting irrelevant responses. His child learns to speak, the spoken instruction shared between the child and adult are taken over by the child, who uses externally stated and often detailed instructions to guide his or her own behavior. By the age of 4 to 4 ½, a trend towards internal and contract speech (inner speech) gradually appears. The child begins to regulate and subordinate his behavior according to his speech. Speech, in addition to serving communication thought, becomes a major self-regulatory force, creating systems of connections for organizing active behavior inhibiting actions irrelevant to the task at hand. (p. 34)

Culture influences the development of higher cognitive functioning through a variety of different channels. Luria (1979) emphasized the importance of the frontal lobes in language, organization, and direction of behavior and speech as a cultural tool that furthers the development of the anterior brain region and self-regulation. Cultural experiences accelerate the use of planning and self-regulation and the other cognitive processes. Luria (1979) suggested that abstraction and generalizations are themselves products of the cultural environment. Children learn, for example, to selectively pay attention to items that are pertinent through conversations and playful interactions with adults. Even simultaneous and successive processes are influenced by cultural experiences (e.g., learning dances, poems, game rules, and so on). Naglieri (2003) summarized research that showed that the influence of social interaction on children's use of plans and strategies resulted in improvements in performance on academic tasks. Luria's concept of functional units and their relationship to the larger sociocultural context provides the foundation for the PASS theory.

From Luria to PASS Theory of Intelligence

The four processes in the PASS theory represent a fusion of cognitive and neuropsychological constructs including executive functioning (planning); selective, sustained, and shifting attention (attention); visual-spatial tasks (simultaneous); and serial features of language and memory (successive) (Naglieri and Das 2005). These four processes are more fully described in the sections that follow.

The human ability to plan differentiates humans from other primates. Planning is associated with the prefrontal cortex. The prefrontal cortex "plays a central role in forming goals and objectives and then in devising plans of action required to attain these goals. The cognitive processes required to implement plans, coordinate these activities, and apply them in a correct order are subserved by the prefrontal cortex. Finally, the prefrontal cortex is responsible for evaluating our actions as success or failure relative to our intentions" (Goldberg 2009, p. 23). Planning helps one to achieve goals through the development of strategies necessary to accomplish tasks for which a solution is required. Therefore, planning is an essential ability to all activities that demand the child or adult to figure out how to solve a problem. This includes self-monitoring and impulse control as well as making, assessment, and implementation of a plan. Thus, planning allows for the generation of solutions, discriminating use of knowledge and skills, as well as control of attention, simultaneous, and successive processes (Das et al. 1996).

The essential dimension of the construct of planning as defined by Naglieri and Das (1997b) is very similar to the description of executive function provided by others (see Naglieri and Goldstein 2006). For example, O'Shanick and O'Shanick (1994) describe executive functions as including the abilities to formulate and set goals, assess strengths and weaknesses, plan and/or direct activities, initiate and/or inhibit behavior, monitor current activities, and evaluate results. Executive functions include abilities to formulate a goal, to plan, to carry out goal-directed behaviors effectively, and to monitor and self-correct spontaneously and reliably (Lezak et al. 2012). McCloskey et al. (2009) identify two key dimensions that unify several diverse definitions of executive functions. To some degree, all definitions address components that direct and cue other processes, and all address functions that link activation to the frontal lobe regions. These skills are essential for fulfilling most daily responsibilities and maintaining appropriate social behavior. A variety of assessment tools that have been proposed to assess executive functions often yield conflicting data given the very broad definition of these functions (e.g., for a review of this issue in the assessment of ADHD, see Barkley 2006). Planning in the PASS theory offers a more finite description that may be characterized as executive function.

Attention is a cognitive process that is closely connected to the orienting response. Attention, as ability, allows a person to demonstrate focused, selective cognition over time, with resistance to distraction. Attention occurs when a person selectively focuses on particular stimuli and inhibits responses to competing stimuli. The process is involved when we must demonstrate focused, selective, sustained, and effortful activity. *Focused* attention involves concentration directed toward a particular activity, and *selective* attention is important for the inhibition of responses to distracting stimuli. *Sustained* attention refers to the variation of performance over time, which can be influenced by the varying amounts of effort required to solve a task. Brain structures within Luria's first functional unit, the reticular formation, allow one to focus selective attention toward a stimulus over a period of time without the loss of attention to other competing stimuli. The longer attention is needed, the more the activity necessitates vigilance. Intentions and goals mandated by the planning process control attention, while knowledge and skills play an integral part in the process as well. The attention work of Schneider et al. (1984) and the attention selectivity work of Posner and Boies (1971), which relates to deliberate discrimination between stimuli, are similar to the way that the attention process is conceptualized. Planning processes regulate a variety of other processes, including attention.

Simultaneous processing is necessary for synthesizing separate elements into a cohesive whole or interrelated group. The ability to recognize patterns as interrelated elements is made possible by the parieto-occipital-temporal brain regions. Due to the substantial spatial characteristics of most simultaneous tasks, there is a visual-spatial dimension to activities that demand this type of process. Conceptually, the examination of simultaneous processing is achieved using tasks that could be described as involving visual-spatial reasoning, found in progressive matrices tests like those developed by Penrose and Raven (1936). Simultaneous processing is not, however, limited to nonverbal content, as demonstrated by the important role it plays in the grammatical components of language and comprehension of word relationships, prepositions, and inflections (Naglieri 1999). This is most apparent in the inclusion of the verbal-spatial relationship subtest in the CAS (Naglieri and Das 1997a). Typically, however, matrices tests have been included in the so-called nonverbal scales of intelligence tests such as the *Wechsler Nonverbal Scale of Ability* (Wechsler and Naglieri 2006), the perceptual reasoning portion of the *Wechsler Intelligence Sale for Children-IV* (WISC-IV; Wechsler 2003), the *Stanford-Binet Fifth Edition* (SB5; Roid 2003), the *Naglieri Nonverbal Ability Test* (NNAT; Naglieri 1997), and the *Kaufman Assessment Battery for Children, Second Edition* (KABC-2; Kaufman and Kaufman 2004) and as a simultaneous processing test (Naglieri and Das 1997a, b).

Successive processing is relevant when working with stimuli arranged in a defined serial order such as remembering or completing information in compliance with a specific order. Successive processing is typically an essential element involved with the serial organization of sounds, such as learning sounds in sequence and early reading. Furthermore, successive processing has been conceptually and experimentally related to the concept of phonological analysis (Das et al. 1994b). When serial information is grouped into a pattern, however, (like the number 553669 organized into 55-3-66-9), then successful repetition of the string may be a function of another cognitive processes, such as planning (i.e., using the strategy of chunking) and simultaneous (organizing the numbers into related groups). This method is often used by older children and can be an effective strategy for those who are weak in successive processing (see Naglieri and Pickering 2003). In clinical practice, we have observed that young children with poor successive processing often have difficulty following directions or comprehending what is being said to them when sentences are too lengthy. Teachers and parents often misinterpret this weakness as a failure to comprehend or as a problem of attention.

Traditionally, intelligence is measured through verbal, nonverbal, and quantitative tests, yet the PASS theory offers an alternative approach to intelligence. This theory broadens the idea of what abilities should be measured and also emphasizes the significance of basic neurocognitive processes. Additionally, the functional units of the brain that encompass the PASS processes are considered the building blocks of ability conceptualized within a neurocognitive processing framework. While the theory may have its roots in neuropsychology, "…its branches are spread over developmental and educational psychology" (Varnhagen and Das 1986, p. 130). Thus, with its connections to developmental and cognitive processing, the PASS theory offers an advantage in explanatory power over the notion of general intelligence (Naglieri and Das 2002).

Measuring PASS Processes

The PASS theory was operationalized by the CAS (Naglieri and Das 1997a; Naglieri et al. 2014a, b). This instrument is thoroughly described in the CAS-2 interpretive handbook (Naglieri et al. 2014b). Naglieri et al. (2014a, b) generated tests to measure the PASS theory following a systematic and empirically based test development program designed to obtain efficient measures of the processes for individual administration. The PASS theory was used as the foundation of the CAS, so the *content of the test was determined by the theory* and not influenced by previous views of ability. The CAS-II (Naglieri et al. 2014a, b) is a Lurian-based test of cognitive abilities and processing that is highly predictive of academic learning and very useful in identifying processing strengths and weaknesses. The CAS-II, normed for ages 5–18, is supported by research linking the CAS-II processes with specific types of learning (Naglieri and Das 1997a, b) and research linking specific-processing deficits with specific learning disabilities. The four CAS-II composites are intended to measure planning, attention, simultaneous processing, and successive processing. The CAS-II composites and all of the subtests can be categorized as measures of cognitive processing. The CAS-II does not contain any measures of verbal knowledge or crystallized intelligence (see Chapter XXX for a thorough presentation of the CAS-II).

Validity

Using Luria's neuropsychological framework of three functional units, Das (1972) and Das et al. (1975, 1979, 1994b) began the task of figuring out methods for measuring the PASS processes. These efforts included extensive analysis of the methods used by Luria, related procedures used within neuropsychology, experimental research in cognitive and educational psychology, and related areas. Their work was summarized in several books by Kirby (1984); Kirby and Williams (1991); Das et al. (1994b); Naglieri (1999); and Naglieri et al. (2014a, b), which provide considerable evidence that the PASS processes associated with Luria's concept of the three functional units could be measured and that once measured, these processes have considerable reliability and validity. Their work also demonstrated that there was significant potential for the application of the theoretical conceptualization of basic psychological processes. The remainder of this section will provide a summary of relevant validity research on the PASS theory as operationalized by the CAS.

Relationship to Achievement

One of the purposes of an ability test is to determine a child's level of cognitive functioning that can then be used to anticipate performance in a number of contexts, such as school. Some have noted that the relationship between a test of ability and achievement is perhaps one of the most important aspects of validity (Brody 1992; Cohen et al. 1992; Naglieri and Bornstein 2003). For many years, researchers have studied the relationship between ability and achievement. Well-known IQ tests often include measures of vocabulary, general information, and arithmetic, as do tests of achievement. It is no surprise then that the relationship between ability and intelligence has been found to be about 0.55–0.60 (Brody 1992; Naglieri 1999). It has been argued, however, that a portion of the correlation between traditional IQ tests and academic achievement tests is due to the similarity in content that exists between these two types of tests (Naglieri and Bornstein 2003; Naglieri and Rojahn 2004). Given that the CAS does not include test items that are typically part of traditional IQ tests such as vocabulary and arithmetic, how well does it correlate with achievement?

Naglieri and Rojahn (2004) studied the relationship between the PASS processing scores of the CAS with the *Woodcock Johnson-Revised Tests of Achievement* (WJ-R; Woodcock and Johnson 1989) with a sample of 1,559 students aged 5–17 years. The correlation between the CAS Full Scale and the WJ-R Tests of Achievement was 0.71 for the standard (all 12 subtests) and 0.70 for the basic battery score (eight subtests). These findings provide evidence for the construct validity of the CAS and more importantly suggest that basic psychological processes are strongly related to academic performance as measured by this standardized test of achievement.

Naglieri et al. (2006) compared the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler 1991) to the CAS and the Woodcock-Johnson III Tests of Achievement (WJ-III-ACH; Woodcock et al. 2001) with a sample of children aged 6–16 who were referred for evaluation due to learning problems. The correlation of the WJ-III-ACH scores with the WISC-III Full Scale IQ scores was 0.63 and 0.83 with the CAS Full Scale. However, the CAS Full Scale scores correlations were significantly higher (Naglieri et al. 2006).

The findings provide evidence for the construct validity of the CAS and suggest that basic psychological processes are strongly correlated with academic performance and are especially important because the measures of the PASS processes do not include achievement-like subtests (e.g., vocabulary and arithmetic). This provides considerable advantage and is especially important for children who come from disadvantaged environments as well as those who have had a history of academic failure.

Importantly, Naglieri and Rojahn (2004) also found that prediction of achievement was slightly higher for the four PASS Scales than the CAS Full Scale. These findings suggested that the four PASS Scales individually and collectively correlate higher with achievement than the four scales aggregated into the one Full Scale score. Additionally, the predictive power of the combination of the four PASS Scales was weakened when any one of the PASS Scales was excluded in the prediction equation (Naglieri and Rojahn 2004). This suggests that each of the PASS Scales has additive value in predicting achievement and further supports the notion of interrelated neurocognitive processes within the Luria's framework of functional units.

Relationship to Behavior

Limited research has been conducted specifically examining the relationship of PASS processes to behavior. Clinically, the connection between PASS processes and a child's behavior is often observed. For example, successive processing involving the ability to follow information in a linear organization or chainlike progression will exert a significant impact on a child's behavior. Planning processing involves the ability to focus one's thinking, attend and screen out distractions which is essential for children to play effectively with others on the playground, and interact with adults, as well as a variety of real-life tasks.

Several researchers have examined the relationship between the behavioral difficulties seen in children with ADHD and PASS profile scores. For example, Paolitto (1999) studied matched samples of ADHD and normal children. Children with ADHD earned significantly lower scores on the planning scale. Similarly, Dehn (2000) and Naglieri et al. (2003) found that groups of children who met diagnostic criteria for ADHD earned significantly lower mean scores on the planning scale of the CAS. These results support the view that ADHD involves problems with behavioral inhibition and self-control, which is associated with poor executive control (Planning; Naglieri and Goldstein 2006). These findings suggest that the PASS processing theory has utility for differential diagnosis, intervention, as well as response to intervention for behavioral problems (Naglieri 2003, 2005).

The measurement of cognitive processes using the Cognitive Assessment System (CAS) has been utilized with individuals who suffer from traumatic brain injury (TBI). Due to the fact that cognitive impairments and deficits are very common in individuals with TBI, the CAS is a measure that can be used to assess the cognitive processes of this population. According to Luria (1973), when one suffers severe brain damage, it is likely that he or she will also experience impairments in such processes as organization and planning (as cited in Gutentag et al. 1998). One of the main reasons as to why an assessment tool such as the CAS would be particularly useful for the TBI population is because typical intelligence tests only yield results that reflect one's general intelligence; they do not provide measurement of basic psychological processes. For example, deficits in attention and planning that interfere with the academic performance of children with TBI (Savage and Wolcott 1994) must be measured. Gutentag et al. (1998) studied children with TBI showed deficits on the CAS compared to a matched control group drawn from the CAS normative population. Neurocognitive deficits were most pronounced in the attention and planning domains and less severe in the simultaneous and successive domains (Gutentag et al. 1998, p. 265).

Fairness

The CAS has been increasingly used across cultures and languages. In Spain for example, Perez-Alvarez et al. (2006) used the CAS in a study assessing the effects of topiramato (a pharmacological treatment for epilepsy) on cognitive processes and behavior. The 35 patients ranging in age from 5 to 15 years were assessed with the CAS at baseline and again at 6 and 12 months. The parents were given behavior-rating scales at each interval as well. At baseline, 6 and 12 months, patients had lower successive scores. At 12 months, planning scores had increased significantly, while there was a concomitant improvement on behavior as measured by rating scales. Mccrea (2009) studied three patients with unilateral focalized stroke lesions longitudinally on the CAS subtests at 1 month and 6 months post infarct, such that each patient functioned as their own baseline. Patient 1 with a left temporal pole lesion had a severe syntactic comprehension deficit on sentence questions. Patient 2 had a rare right anterior cerebral artery (ACA) aneurysm culminating in an orbitofrontal syndrome and impairments on expressive attention, word series, as well as a praxis-based figure ground reversal phenomenon on figure memory. Patient 3 suffered a right frontoparietal lesion with resulting representational as well as elements of motor neglect and impairments on matching numbers, number detection, and receptive attention. Each patient's lesions were all entirely consistent with the nature of cognitive neuropsychological symptoms suggesting that the CAS subtests are unique and also sensitive and specific to focalized cortical lesions.

The characteristics of the US population continue to change with every census, and the need for fair assessment of children has become progressively more important. Traditional IQ tests have items that measure content that is dependent on exposure to the dominant culture, language, and formal education. This content can create an unfair disadvantage for many children, such as those living in non-English-speaking homes and impoverished environments. Reducing the amount of knowledge needed to correctly answer the questions on intelligence tests is a useful way to ensure appropriate and fair assessment of diverse populations. Some researchers have suggested that conceptualizing intelligence as a set of psychological processes, such as the PASS theory as operationalized by the CAS, has utility for assessment of children from culturally and linguistically diverse populations because verbal and quantitative skills are not included (Naglieri and Otero 2011; Naglieri et al. 2005, 2007).

Several researchers have found up to a 15-point mean difference between Blacks and Whites on traditional tests of cognitive ability.

Results for PASS processing tests have shown only small differences between these groups. For example, Naglieri et al. (2005) compared CAS scores of a sample composed of 298 Black children and 1,691 White children. Controlling for key demographic variables, regression analyses showed an estimated CAS Full Scale mean score difference of 4.8, which is smaller than that found with traditional tests of ability. Another finding was that correlations between the CAS scores and WJ-R Tests of Achievement were very similar for Blacks (0.70) and Whites (0.64; Naglieri et al. 2005). Naglieri et al. (2006) examined CAS scores for 244 Hispanic and 1,956 non-Hispanic children. They found that the two groups differed by 6.1 points when the samples were unmatched samples, 5.1 with samples matched on basic demographic variables, and 4.8 points when demographic differences were statistically controlled. These findings further demonstrate the utility of PASS theory as one way to fairly assess diverse populations.

When evaluating the cognitive ability of English language learners, psychologists currently have three practice options (Ortiz 2009). These methods are (a) modifications and adaptations of the standardized administration and scoring of the test; (b) the selection and use of specific tests or battery of tests that are of a nonverbal nature; and (c) the use of a more traditional native language-based test (e.g., WISC-IV Spanish). Each method has its limitations and advantages. The first method, modifying or adapting the tests in terms of administration or scoring, violates standardization directly, resulting in error and in less reliability and validity of scores attained. The second method, and perhaps the most commonly practiced (Ortiz 2009), involves the use of a nonverbal battery or the administration of select subtests that make up the PIQ (Figueroa 1990), or more recently, the PRI. This method, although it reduces the impact of language on test results, would not be helpful in cases in which the student's dysfunction is actually language based (such as reading or written language). The third option, the use of a native language test, may seem to be the ideal option with ELL students. However, these tests fail to control for the level of language proficiency (Harris and Llorente 2005).

In the case of ELL Hispanic children, Naglieri et al. (2007) compared the English and Spanish versions of the CAS for bilingual Hispanic children. The children in this study earned very similar CAS Full Scale scores, and deficits in successive processing were found on both versions of the test. Importantly, 90 % of children who had a neurocognitive weakness on one version of the CAS also had the *same* neurocognitive weakness on the other version of the CAS. Otero et al. (2013) examined the performance of referred Hispanic English Language Learners of varying proficiency levels ($N=40$) on the English and Spanish versions of the *CAS* and found no significant differences between the Full Scale scores or in any of the PASS scales. Students earned their lowest scores in successive processing regardless of the language in which the test was administered. These findings suggest that the *CAS* may be a useful measure for Hispanic children with underdeveloped English language proficiency. These results suggest that the PASS scores from both the English and Spanish version of the CAS could be used as part of a comprehensive evaluation.

Conclusions

Although several definitions of intelligence have been set forth for over a century, the definition of intelligence as a constellation of neurocognitive processes such as planning, attention, simultaneous, and successive processes is unique and theory driven. There is a growing need for neurocognitive measures to evaluate and explain function, facilitate prognosis, and most importantly guide intervention. Luria's PASS theory offers a blueprint for defining the basic neurocognitive processes underlying human performance, behavior, and intelligence. Appreciation and application of this processing model as a framework for assessment provides psychologists with an essential tool necessary to not just understand children's learning and behavior but to guide and develop effective intervention. The PASS theory as operationalized by the CAS-2 provides a well-developed tool for assessment of the four basic psychological processes described by Luria. The importance of assessing PASS neurocognitive processes cannot be overemphasized in light of the evidence of its use with various groups, including Hispanic English language learners.

References

Bostan, A. C., & Strick, P. L. (2010). The cerebellum and basal ganglia are interconnected. *Neuropsychology Review, 20*, 261–270.

Bostan, A. C., Dum, R. P., & Strick, P. L. (2010). The basal ganglia communicate with the cerebellum. In: *Proceedings of the national academy of sciences* (Vol. 107, pp. 8452–8456).

Bostan, A. C., Dum, R. P., & Strick, P. L. (2013). Cerebellar networks with the cerebral cortex and basal ganglia. *Trends in Cognitive Sciences, 17*, 241–254.

Brody, N. (1992). *Intelligence*. San Diego: Academic.

Cohen, R. J., Swerdlik, M. E., & Smith, D. K. (1992). *Psychological testing and assessment*. Mountain View: Mayfield Publishing.

Das, J. P. (1972). Patterns of cognitive ability in nonretarded and retarded children. *American Journal of Mental Deficiency, 77*, 6–12.

Das, J. P. (1980). Planning: Theoretical considerations and empirical evidence. *Psychological Research (W. Germany), 41*, 141–151.

Das, J. P., Kirby, J. R., & Jarman, R. F. (1975). Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin, 82*, 87–103.

Das, J. P., Kirby, J. R., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processes*. New York: Academic.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes*. Needham Heights: Allyn & Bacon.

Das, J. P., Kar, B. C., & Parrila, R. K. (1996). *Cognitive planning: The psychological basis of intelligent behavior*. Thousand Oaks: Sage.

Das, J. P., Parrila, R. K., & Papadopoulos, T. C. (2000). Cognitive education and reading disability. In A. Kozulin & Y. Rand (Eds.), *Experience of mediated learning* (pp. 276–291). Amsterdam: Pergamon.

Dehn, M. J. (2000). *Cognitive assessment system performance of ADHD children.* Paper presented at the annual NASP Convention, New Orleans.

Figueroa, R. A. (1990). Assessment of linguistic minority group children. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Volume 1 intelligence and achievement* (pp. 135–152). New York: Guilford.

Friston, K. (2002). Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annual Review of Neuroscience*, 25, 221–250.

Goldberg, E. (2009). *The new executive brain: Frontal lobes and the civilized mind*. New York: Oxford University Press.

Gutentag, S. S., Naglieri, J. A., & Yeates, K. O. (1998). Performance of children with traumatic brain injury on the cognitive assessment system. *Assessment, 5*, 263–272.

Harris, J. G., & Llorente, A. M. (2005). Cultural considerations in the use of the Wechsler Intelligence Scale for Children-Fourth Edition. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 382–416). San Diego: Academic.

Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., de Haan, M., Tucker, L. A., Baron-Cohen, S., & Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology, 17*, 599–619.

Kar, B. C., Dash, U. N., Das, J. P., & Carlson, J. S. (1992). Two experiments on the dynamic assessment of planning. *Learning and Individual Differences, 5*, 13–29.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children* (2nd ed.). Circle Pines: American Guidance Service.

Kirby, J. R. (1984). *Cognitive strategies and educational performance*. New York: Academic.

Kirby, J. R., & Williams, N. H. (1991). *Learning problems: A cognitive approach*. Toronto: Kagan and Woo.

Kolb, B., & Whishaw, I. Q. (2009). *Fundamentals of human neuropsychology* (6th ed.). New York: Worth Publishers.

Kolb, B., Gibb, R., & Robinson, T. E. (2003). Brain plasticity and behavior. *Current Directions in Psychological Science, 12*, 1–4.

Korkman, M. (1999). Applying Luria's diagnostic principals in the neuropsychological assessment of children. *Neuropsychology Review, 9*, 89–105.

Koziol, L. F., & Budding, D. E. (2009). Subcortical structures and cognition : implications for neuropsychological assessment. New York, NY: Springer.

Koziol, L. F., & Stevens, M. C. (2012). Neuropsychological assessment and the paradox of ADHD. Applied Neuropsychology: Child, 1(2), 79-89.

Koziol, L., Joyce, A.W., & Wurglitz, G. (2014). The neuropsychology of attention: Revisiting the "Mirsky Model". Applied Neuropsychology: Child. 3(4), 297-307.

Koziol, L. F. (2014). *The myth of executive functioning: Missing elements in conceptualization, evaluation and assessment*. New York: Springer.

Koziol, L. F., Budding, D. E., & Chidekel, D. (2011). Sensory integration, sensory processing, and sensory modulation disorders: Putative functional neuroanatomic underpinnings. *Cerebellum, 10*, 770–792.

Lewandowski, L., & Scott, D. (2008). Introduction to neuropathology and brain-behavior relationships. In R. C. D'Amato & L. C. Hartlage (Eds.), *Essentials of neuropsychological assessment: Treatment planning for rehabilitation* (2nd ed.). New York: Springer Publishing Company.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, T. (2012). *Neuropsychological assessment* (5th ed.). New York: Oxford University Press.

Luria, A. R. (1962) *Higher Cortical Functions in Man*. Moscow University Press. Library of Congress Number: 65-11340.

Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper and Row.

Luria, A. R. (1973). *The working brain*. New York: Basic Books.

Luria, A. R. (1979). *The making of mind: A personal account of Soviet psychology*. Cambridge, MA: Harvard University Press.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.

Luria, A. R. (1982). *Language and cognition*. New York: Wiley.

McCloskey, G., Perkins, L. A., & Van Diver, B. (2009). *Assessment and intervention for executive function difficulties*. New York: Routledge.

Mccrea, S. M. (2009). A cognitive neuropsychological examination of the Das-Naglieri cognitive assessment system subtests: A report of three stroke cases studied longitudinally during recovery. *International Journal of Neuroscience, 119*(4), 553–559.

Mirsky, A. F., & Duncan, C. C. (2001). A nosology of disorders of attention. *Annals of New York Academy of Science, 931*, 17–32.

Mirsky, A. F., & Duncan, C. C. (2003). Attention battery for children: A systematic approach to assessment. In G. Goldstein & S. R. Beers (Eds.), *Comprehensive handbook of psychological assessment, volume 1, intellectual and neuropsychological assessment*. New York: Wiley.

Naglieri, J. A. (1997). *Naglieri nonverbal ability test*. San Antonio: The Psychological Corporation.

Naglieri, J. A. (2003). Current advances in assessment and intervention for children with learning disabilities. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities Volume 16: Identification and assessment* (pp. 163–190). New York: JAI.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21*, 244–260.

Naglieri, J. A., & Conway, C. (2009). The cognitive assessment system. In J. A. Naglieri & S. Goldstein (Eds.), *A practitioner's guide to assessment of intelligence and achievement* (pp. 3–10). New York: Wiley.

Naglieri, J. A., & Das, J. P. (1997a). *Cognitive assessment system*. Itasca: Riverside Publishing Company.

Naglieri, J. A., & Das, J. P. (1997b). *Cognitive assessment system interpretive handbook*. Chicago: Riverside Publishing Company.

Naglieri, J. A., & Das, J. P. (2002). Practical implications of general intelligence and PASS cognitive processes. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 855–884). New York: Erlbaum.

Naglieri, J. A., & Das, J. P. (2005). Planning, attention, simultaneous, successive (PASS) theory: A revision of the concept of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (2nd ed., pp. 136–182). New York: Guilford.

Naglieri, J. A., & Goldstein, S. (2006). The role of intellectual processes in the DSM-V diagnosis of ADHD. *Journal of Attention Disorders, 10*, 3–8.

Naglieri, J. A., & Otero, T. (2011). Cognitive Assessment System: Redefining intelligence from a neuropsychological perspective. In A. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 320–333). New York: Springer.

Naglieri, J. A., & Otero, T.M. (2012). The Cognitive Assessment System. In D. P. Flanagan and P. L. Harrison (Eds.) *Contemporary Intellectual Assessment* (Third Edition) (pp. 379-399). New York: Guilford.

Naglieri, J. A., & Pickering, E. (2003). *Helping children learn: Instructional handouts for use in school and at home*. Baltimore: Brookes.

Naglieri, J. A., & Rojahn, J. R. (2004). Validity of the PASS theory and CAS: Correlations with achievement. *Journal of Educational Psychology, 96*, 174–181.

Naglieri, J. A., Goldstein, S., Iseman, J. S., & Schwebach, A. (2003). Performance of children with attention deficit hyperactivity disorder and anxiety/depression on the WISC-III and Cognitive Assessment System (CAS). *Journal of Psychoeducational Assessment, 21*, 32–42.

Naglieri, J. A., Salter, C. J., & Edwards, G. (2004). Assessment of ADHD and reading disabilities using the PASS theory and cognitive assessment system. *Journal of Psychoeducational Assessment, 22*, 93–105.

Naglieri, J. A., Rojahn, J. R., Matto, H. C., & Aquilino, S. A. (2005). Black white differences in intelligence: A study of the PASS theory and cognitive assessment system. *Journal of Psychoeducational Assessment, 23*, 146–160.

Naglieri, J. A., Rojahn, J., & Matto, H. (2007). Hispanic and non-Hispanic children's performance on PASS cognitive processes and achievement. Intelligence, 35, 568–579.

Naglieri, J. A., Goldstein, S., Delauder, B. Y., & Schwebach, A. (2006). WISC-III and CAS: Which correlates higher with achievement for a clinical sample? *School Psychology Quarterly, 21*, 62–76.

Naglieri, J. A., Rojahn, J., & Matto, H. (2007). Hispanic and non-Hispanic children's performance on PASS cognitive processes and achievement. *Intelligence, 35*, 568–579.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014a). *Cognitive assessment system* (2nd ed.). Austin: Pro-Ed Publishing Company.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014b). *Cognitive assessment system-second edition. Interpretive handbook*. Austin: Pro-Ed Publishing Company.

Neville, H. (2006). Different profiles of plasticity within human cognition. In Y. Munakata & M. Johnson (Eds.), *Processes of change in brain and cognitive development*. New York: Oxford University Press.

Neville, H., & Stevens, C. (2008). Experience shapes human brain development and function: A framework for planning interventions for children at-risk for school failure (Summary). Annual meeting of the American Association for the Advancement of Science, Boston.

O'Shanick, G. J., & O'Shanick, A. M. (1994). Personality and intellectual changes. In J. M. Silver, S. C. Yudofsky, & R. E. Hales (Eds.), *Neuropsychiatry of traumatic brain injury* (pp. 163–188). Washington, DC: American Psychiatric Press.

Ortiz, S. O. (2009). Bilingual-multicultural assessment with the WISC-IV. In E. O. Lichtenberger & A. S. Kaufman (Eds.), *Essentials of the WISC-IV* (pp. 295–309). New York: Wiley.

Otero, T. M., Gonzalez, L., & Nagleri, J. A. (2013). The neurocognitive assessment of Hispanic English-language learners with reading failure. *Applied Neuropsychology: Child, 2*(1), 24–32.

Paolitto, A. W. (1999). Clinical validation of the Cognitive Assessment System with children with ADHD. *ADHD Report, 7*, 1–5.

Penrose, L. S., & Raven, J. C. (1936). A new series of perceptual tests: Preliminary communication. *British Journal of Medical Psychology, 16*, 97–104.

Perez-Alvarez, P., Timoneda-Gallart, C., & Baus-Rosell, J. (2006). Topiramato y epilepsia a la luz del Das-Naglieri Cognitive Assessment System. *Revista Neurologia, 42*(1), 3–7.

Pluker, J. A., & Esping, A. (2014). *Intelligence 101*. New York: Springer.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*, 391–408.

Roid, G. (2003). *Stanford-Binet fifth edition*. Itasca: Riverside.

Savage, R. C., & Wolcott, G.F. (Eds). (1994). Educational dimensions of acquired brain injury. Austin, TX: PRO-ED

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and controlled processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 1–28). New York: Academic.

Semrud-Clikeman, M., & Teeter Ellison, P. A. (2009). *Child neuropsychology: Assessment and intervention of neurodevelopmental disorders* (2nd ed.). New Jersey: Springer.

Stuss, D. T., & Benson, D. F. (1990). The frontal lobes and language. In E. Goldberg (Ed.), *Contemporary psychology and the legacy of Luria*. Hillsdale: Lawrence Erlbaum Associates.

Thorndike, R. L. (1997). *Measurement and evaluation in psychology and education* (6th ed.). Upper Saddle River: Prentice Hall.

Varnhagen, C. K., & Das, J. P. (1986). Neuropsychological functioning and cognitive processing. In J. E. Obzrut & G. W. Hynd (Eds.), *Child neuropsychology, Vol. 1: Theory and research* (pp. 117–140). New York: Academic.

Vygotsky, L. S. (1976). *Mind in society: Development of higher psychological processes*. London: Harvard University Press.

Wechsler, D. (1991). *Wechsler intelligence scale for children* (3rd ed.). San Antonio: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio: The Psychological Corporation.

Wechsler, D., & Naglieri, J. A. (2006). *Wechsler nonverbal scale of ability*. San Antonio: Harcourt Assessment.

Wilkinson, A., & Sermund-Clikeman, M. (2008). *Motor speed in children and adolescents with nonverbal learning disabilities.* Paper presented at the International Neuropsychological Society, Waikoloa, Hawaii.

Woodcock, R. W., & Johnson, M. B. (1989). *WJ-R Tests of Cognitive Ability*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III. Itasca, IL: Riverside Publishing.

CHC Theory of Intelligence

15

Samuel O. Ortiz

The concept of intelligence is as ubiquitous as any other artifact of our culture. It is a word that is used often, readily accepted in meaning, and used to invoke explanations for all sorts of social and behavioral phenomena observed and encountered on a daily basis. It is found frequently in all forms of popular and professional literature and discourse and is a common topic in introductory science texts across myriad specialties and related disciplines, particularly psychology. Even the term born from the idea of intelligence, "IQ" (or "intelligence quotient"), is no longer an acronym per se but an accepted word in all modern language dictionaries including the one that even governs the official rules of the game, Scrabble. Unquestionably, we can point to few things in modern life that can claim to be so familiar, so well known, and so much a part of our daily experience as the notion of intelligence and our presumed understanding of it.

That intelligence is so ingrained in what we know value may also have done it a considerable disservice. Consider, for example, we all readily admit that we know exactly what it refers to, and thus rarely is anyone called upon to define it or its usage in a given context. Even less frequently is someone asked to defend the theoretical or empirical bases that would support its use in a given context. But when a situation requires a precise and scientifically validated definition, it is alarming to see both lay people and professionals scramble for support like rats on a sinking ship. Indeed, perhaps the most natural reaction in such cases is to invoke the one explanation provided by McNemar (1964) which is sure to appease everyone, where he noted "no definition is required because all intelligent people know what intelligence is—it is the thing the other guy lacks" (p. 871).

The focus of this chapter is not centered directly on intelligence. The purpose here is to describe the history and development of what is currently referred to as CHC theory, an acronym derived from the chronological order of contribution by the surname of its three main developers (i.e., Cattell, Horn, and Carroll). It is impossible, however, to discuss the development of the theory without addressing the debates surrounding the very notion and definition of intelligence, since any such theory must necessarily provide an explanation of it. And despite offering what is considered perhaps the most empirically validated and best supported theory of human cognitive abilities to date (e.g., Schneider and McGrew 2012), CHC theory is unlikely to be the first mention in response to any query regarding conceptualizations of intelligence. Admittedly, there has been an extraordinary and precipitous increase in published and unpublished research regarding CHC theory since 2000 (Schneider and McGrew 2012),

S.O. Ortiz (✉)
St. John's University,
8000 Utopia Pkwy, Jamaica, NY 11439, USA
e-mail: ortizs@stjohns.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_15,

but the reason why CHC theory remains in relative obscurity compared to other, more popular notions of intelligence (e.g., Spearman's *g*, Gardner's 'multiple intelligences') is an important part of its history and extremely relevant to its development. Likewise, because the theory itself carries the very names of its constituent authors, discussion of the one cannot be accomplished without discussion of the other.

CHC Theory and Cattell: Birth and the Early Years

The basic formulations of what would eventually become known as CHC theory began with Raymond B. Cattell (1941, 1943) whose academic pedigree in psychology could be traced back to his dissertation mentor, Charles Spearman, who himself was trained by Wilhelm Wundt. Born in 1905, it is important to recall and understand the time in which Cattell was attending school and how he was influenced heavily not only by his education and academic mentoring but also the values and beliefs of his culture and the prevailing zeitgeist. Having just missed military service in World War I, Cattell entered University College in London in 1921 studying chemistry primarily but then moving into psychology for graduate studies in 1925 at Kings College. Cattell had the good fortune of working with Spearman and completing his dissertation right around the time Spearman was refining the process of evaluating correlational data statistically in a manner that would eventually become known as factor analysis. Needless to say, Cattell was very well trained in this relatively new method and adopted not only the methodology from Spearman but a good deal of his philosophy as well. This included a strong belief in the genetic nature and determination of intelligence, ideas regarding the need for proactive policy with regard to individual's identified with low intelligence, and perhaps the most salient idea regarding his future career—a staunch acceptance of data analytic methods that provided a first-order general factor which was believed to represent *g* or general intelligence. It should be noted that Cattell's beliefs were no different than the vast majority of his fellow colleagues at the time, particularly his British counterparts whose historical class structure was justified on the basis of hereditary power stemming from hereditary attributes—particularly intelligence. Coupled with a society that was still actively engaged in colonialism, albeit wrestling with the notion more so than before and searching for scientific support to justify the practice, it is not surprising that Cattell's own early thoughts concerning intelligence differed little from his mentors or others at the time. What is truly remarkable about Cattell, however, was his willingness to modify his position and accept the mistakes and failings that characterized some of his early beliefs and theoretical conceptions when the data suggested otherwise. For example, in 1997 when Cattell was nominated and selected for the prestigious APA Gold Medal for Lifetime Achievement in the Science of Psychology, he was quickly accused of being a racist and condemned by a small faction who hinged their attack on Cattell's early thoughts on eugenics and theology as well as statements he made some 60 years prior and without any recognition of the manner in which his views evolved over the course of several decades. In an open letter to APA, he declined the award and, perhaps irrevocably heartbroken by the unfair portrayal, died several months later.

Whereas the label of racist might be fairly applied to some psychologists whose formative years spanned the early 1920s to late 1940s, it is not accurate to do the same with Cattell. First, he was a product of the times, and many of those with similar beliefs and attitudes (e.g., C. C. Brigham, Lewis Terman, H. H. Goddard) eventually recanted those beliefs formally in the face of better science, a mature perspective of life, or an intolerance for the ruse that drives various aspects of scientifically based racism. In fact, Cattell not only threw over many of the positions which had become socially undesirable and untenable, he actually departed in a significant way from the very idea used to support such discriminatory policies—the belief that there was one and only one type of intelligence (i.e., general intelligence or *g*) and that it was supreme among all other so-called second-order factors that might be identified. That Cattell was able to break with what is the most important empirical foundation for scientifically driven racist beliefs not only indicates that he was not afraid to relinquish such inequitable ideas as fallacious but that he was also willing to challenge the very core of the prevailing statistical ideology that he had been taught by none other than Spearman himself.

Cattell's application of factor analytic techniques and his early dissatisfaction with the idea of general intelligence or a single *g* began with a very brief presentation at the APA annual conference in 1941 published in the *Psychological Bulletin* that may well have been prompted by the renewed interest in the testing of military recruits as the USA entered World War II. In that paper (Cattell 1941), he stated, "the notion of a general factor does not recommend itself so strongly in adult testing because adult tests are less saturated with a general factor" (p. 592). Two things are noteworthy in his comments. First, that he is beginning to see that *g* is an inadequate explanation for intelligence at the adult level. And second, that intelligence must be viewed as a developmental process. Perhaps one of the greatest and most overlooked contributions to intelligence provided by Cattell (and still reflected in research conducted on CHC theory today) is the attention to the manner in which abilities develop, peak, maintain, or decline across the life span. Cattell concerned himself with developmental issues to an extent that is extremely uncommon even today but which is fundamental to any definition. Cattell's movement away from *g* can also be seen in his admonition that answers to the inadequacies of *g* in adults may rest with the work of Thurstone and Thomson, and he offered a practical solution that forms the very beginnings of what would become known as Gf-Gc theory (one of the precursors to CHC theory). His proposed solution to measurement is better delineated in a later publication (Cattell 1943) where he begins to attack tests on several fronts including inadequate psychometric properties, norm samples lacking representation for the general population, and overreliance on manual dexterity. For the most part, however, he decries the paucity of suitable tests for use in evaluating adults, particularly those over the age of 20, and despite listing some 44 popular tests currently in print, Cattell (1943) denounces the lot as reflecting a "dearth of tests" that "must come as a shock to most psychologists, for it has been widely assumed that the momentum of real progress in intelligence test theory and practice which arose in the first two decades of this century as continued unabated through the ensuing 20 years" (p. 156). That tests have failed to demonstrate any useful progress over the prior two decades is ascribed by Cattell to problems with the fundamental theory of intelligence.

At this point in his career, Cattell had already begun dismissing many of the foundations of the genetic arguments for intelligence that had gone hand in hand with Spearman's *g*. Cattell's focus on adults sensitized him to the problems involved in evaluating intelligence free from the effects of schooling, verbal skills, and mathematical abilities and the need to account for the decline in "speed" but not "power" of intelligence as one ages. Cattell was clearly dissatisfied with notions regarding tests of intelligence and their concomitant lack of theory apart from the three basic rules taken from three different fields of study. Clinical study of individuals with low intelligence provided the notion that intelligence involved the capacity to think abstractly. Animal studies and analogs offered up the idea that intelligence was reflected by an ability to learn. And last, measurement in education generated the premise that intelligence must involve the capacity to adapt means to ends. These three principles were reiterated widely at the time and troubled Cattell greatly as evidenced by his citation of Wechsler (1939) and Wells (1932) to wit: "Wechsler, in his Bellevue Scale manual, defines adult intelligence as 'the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment.' But Wells (p. 265), deploring the fact that Wechsler's concept of intelligence involves adding up the subtest scores into a single total, says: 'The chief use of global scores is administrative.'" (p. 159). Cattell argued forcefully that despite a wide array of published tests of intelligence, few authors attempted to provide positive statements about the nature of intelligence, and those that did were no different than what had been offered even 40 years prior. It was not so much that Cattell simply felt any general theory of intelligence was lacking but more that there was very little support for intelligence irrespective of the theory. He notes that "those applied psychologists who have been the most prolific designers and users of 'intelligence tests' and who have so long and so uncritically accepted the sum of a hodgepodge of tests in the form of a single IQ measurement are now swinging—on no better evidence—to the opposite extreme of demanding that tests should yield measurements of separate abilities" (p. 161). All this is not to say that Cattell was completely disenchanted with the notion of a general factor of intelligence. In discussing his opinions regarding the competing viewpoints of Spearman and Thurstone, Cattell argues that Spearman did admit to finding certain group factors and that Thurstone also conceded that a general factor could be identified. But he leans in favor of Spearman by noting that he "introduces his group factors to the reader with a cold and perfunctory politeness, while Thurstone's general factor is only permitted to enter society as a 'second-order factor' after the 'primary abilities' have made off with all of the actual test variance" (p. 170). Still, Cattell emphasizes that factor analysis will not bring such a debate to an end but that it will permit significantly greater objectivity to the entire enterprise. He understood that disagreements in science will continue to stem from differences in how data are viewed and interpreted, but the arbitrariness of the endeavor is removed and elevates the enterprise to the level that true science demands. Graphical representations of Spearman's and Thurstone's models are provided in Fig. 15.1 for comparative purposes.

Following his own admonitions and concerns about theory, Cattell introduced the notion that deterioration in intelligence in adults was not uniform across all tests but differential, with the higher *g*-saturated tests (e.g., vocabulary, information, verbal comprehension) showing the least decline across the life span and relatively *g*-unsaturated tests (e.g., speeded tasks, abstract problem-solving, unfamiliar performances) showing the most precipitous declines. It must have dawned on him quickly that if there is such a thing as a single general factor, the evidence of abilities that decline or do not decline did not support it. A new conceptualization must be created to explain this phenomenon, and on page 178, he finally writes, for perhaps the first time in formal discourse, that the difference between such abilities can be expressed by the terms "crystallized" when implicating the former and as "fluid" in referring to the latter (Cattell 1943). In describing his "hypothesis of fluid and crystallized ability," (p. 180), Cattell postulates that if sustained, his formulation would be critical to intelligence testing in adults by requiring the specific measurement of both abilities rather than relying on notions of general intelligence. It should be understood as well that Cattell was not wholly dismissing Spearman's *g* or attempting to devalue it and its utility. Rather, he was merely breaking it into two relatively equal pieces—both of which could be considered equivalent components to intelligence that should be viewed as distinct, yet cohesive in a general sense. That he likely hoped to retain ties to his mentor, Spearman, is evident in the data and analyses Cattell uses to support his new theoretical formulation. Cattell took great pains to tease out speed issues, noting that there were various aspects of speed which were related to the decline in fluid abilities. Yet, he did not attempt to incorporate speed as a factor in its own right, either because his analysis of the data did not permit it or he could not reconcile the problems a third factor introduces with respect to the idea that there should be only a single general one. Two parts of the same thing was likely acceptable in principle, and Cattell even used Spearman's designations (i.e., lower case *gf* and *gc*) to maintain the link to a single *g*. But three would likely have proven too much of an anomaly, and the matter remained unaddressed until he mentored a dissertation designed to examine his theory by his student John Horn who had no qualms about identifying the presence of additional factors that had been either conveniently overlooked by Cattell or simply hidden by his own personal and professional biases.

Despite his stated concerns about the glacial pace at which intelligence theory had proceeded

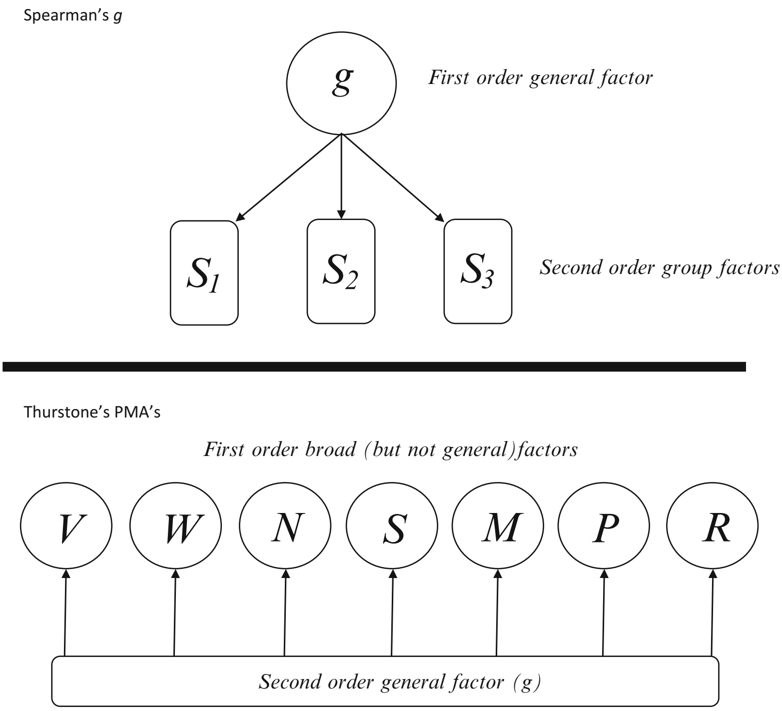


Fig. 15.1 Comparison of Spearman's *g* and Thurstone's primary mental abilities

prior to his explication of a two factor *gf*-*gc* model, Cattell himself allowed two decades of his own to pass before he put his theoretical framework to the test. It was not that he completely neglected the theory, and indeed he offered a few refinements along the way (Cattell 1950; Cattell 1957a, b). But evidence in support of his theory lagged significantly, and Cattell admitted that the bridging of the theory-practice gap was not as simple as it might seem. Part of the problem was likely the industrialization of intelligence tests and testing which had provided an economic boom in its own right and resulted in greater difficulties as well as added costs required for development. Nevertheless, Cattell persisted and after 20 years of theory, he turned his attention to empirical investigation in support of his yet unsubstantiated model. The result was a publication that serves as the official pronouncement of the birth of the "theory of fluid and crystallized intelligence" and outlined a well-developed and modern conceptualization of intelligence that contrasted starkly against the various theoretical frameworks, primarily single *g*, that had been carried over unchanged for the past half century (Cattell 1963). In hindsight, Cattell's inability, reluctance, or other reason for avoiding the necessary research may have also played a role in why his theoretical model was not readily adopted by the major test publishing companies and why, despite the advancements in thinking it accorded, was not used as a platform for the development or revision of current and future intelligence tests. The work of Cattell's student, John Horn, would begin to correct this oversight but not without another three decades of relative obscurity.

CHC Theory and Horn: Headlong into Adolescence

In his landmark paper on *gf*-*gc* theory, Cattell (1963) makes a notable reference to the unpublished work of John Horn, the second major contributor to CHC theory. Cattell emphasized that investigations of *gf*-*gc* theory cannot be made piecemeal or relative to one ability but not the other. He gleefully noted that "the necessary experimental conditions for constructive conclusions [about the theory] are possible only if the total theory is kept in focus, as is happily the case in, for example, the recent work of Horn" (p. 2). Cattell does not cite Horn's research in the bibliography, but the timing of the article suggests that he was most likely referring to Horn's (1965) dissertation which was in progress around that time. During an informal collegial dinner and discussion among a number of so-called CHC enthusiasts, Horn related personal details on the history of his dissertation explaining that he had hoped to do something quite different, and when it did not work out, Cattell provided him with some data and recommended that he analyze it (J. Horn—personal communication, March 1, 2007). He did, and not surprisingly, the title of his thesis is quite on point, "Fluid and crystallized intelligence: A factor analytic study of the structure among primary mental abilities" and indicates that just as Spearman had passed on the legacy of factor analysis to Cattell, so too did Cattell pass on his affinity for factor analysis to Horn, but not necessarily their attachment to a single *g* or to two related *g* factors (*gf*-*gc*).

Horn apparently got a very quick start in his role as a contributor to Cattell's newly formed *gf*-*gc* theory. He had already been studying with Cattell and contributing some research in line with the theory when his dissertation provided a powerful and comprehensive test of the theory, much like Cattell had done in 1963, and inadvertent as his topic may have been, it set the course for much of the rest of his academic career. What Horn found, however, was that Cattell was essentially "wrong" in suggesting that there existed two distinct but similar aspects of general intelligence because his new analyses, completed in the manner specified and approved by Cattell himself, provided support for at least four such primary factors. In addition to what he referred to as "Gf" and "Gc," Horn suggested the data he collected also supported two other primary mental abilities including general visualization (or Gv) and general speediness (or Gs).[1]

Figure 15.2 provides a side-by-side view of Cattell's original *gf*-*gc* formulation and the modifications and conceptual differences suggested by Horn (1965). As is evident in the illustration, Horn retained the basic definitions provided originally by Cattell in terms of Gf and Gc, including the use of the term "intelligence" as part of the name for each one (i.e., fluid intelligence and crystallized intelligence). He viewed these as broad dimensions exhibited as individual differences involving various aspects of reasoning and problem-solving in both, but without any requirement of pre-training, knowledge, or education in the former and as specifically reflected as necessary "skills which constitute the collective intelligence of the culture and which are learned under conditions of intensive acculturation" (p. 309). The construction of Gc in this manner highlights an important dilemma in test construction that seeks to base itself on sound theory. If an ability is defined as "acculturation" and specified to be the "result … of opportunity such as is occasioned by special schooling and continued exposure to the culture in again," the practice of stratifying normative samples primarily on the basis of age becomes problematic. Certainly, education can be easily controlled as it is not difficult to ascertain the extent to which an individual has received formal instruction over the life span. But rarely is that information coupled with actual age in determining an appropriate standard for comparison. More troubling, however, is the lack of attention paid to the process of cultural knowledge acquisition outside of school that Horn points out as being related to age.

[1] It should be noted here that it is not clear why Horn chose to use capital letters in designating these abilities. Whereas it was not an uncommon practice in the literature at the time, it may have been a deliberate choice on Horn's part so as to reflect his desire to eliminate all direct links to any of the abilities being misconstrued as components of a larger general factor, or g—an idea he was beginning to accept as untenable even in his dissertation. Whatever the case, the capital letter designations have become the default and preferred format within CHC theory and are used throughout the remainder of the chapter.

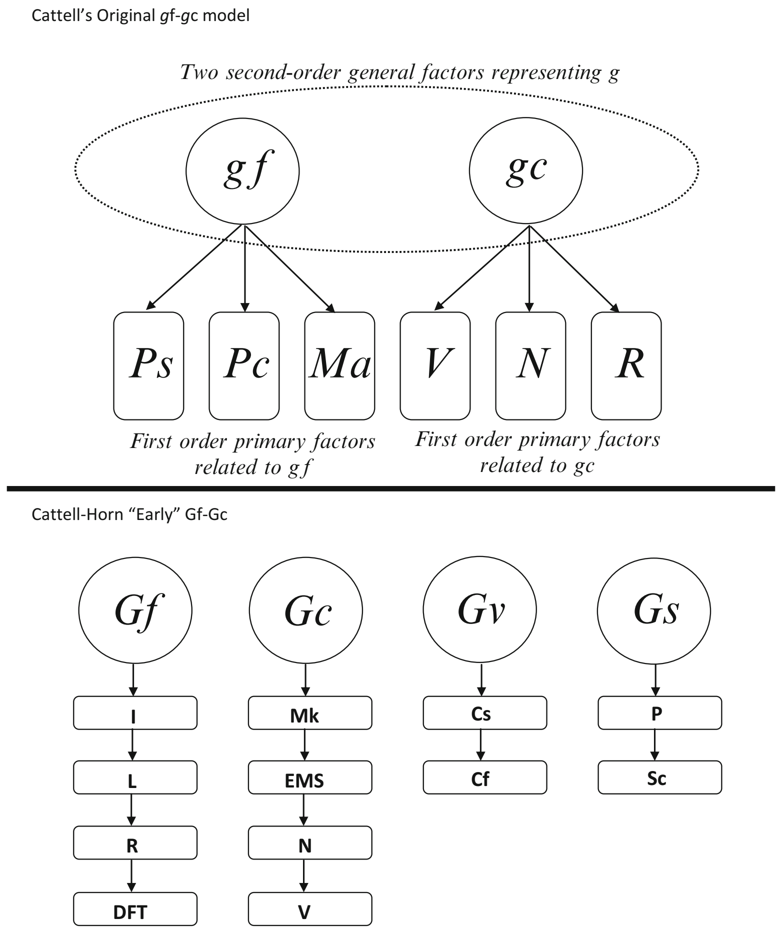


Fig. 15.2 Comparison of Cattell's *gf*-*gc* theory and the Cattell-Horn early Gf-Gc theory

Whereas individuals residing in the same country under relatively similar conditions would be expected to share approximately the same opportunities for learning about the culture in which they live, the same cannot be said for individuals whose opportunities for learning have been limited in comparison to same age peers due to factors such as immigration, ethnic and cultural differences in the home and community, and language differences. Moreover, in what manner can Gc be measured without the use of Gc-based stimuli? As Cattell so proudly announced, the theory of fluid and crystallized abilities must be evaluated as a whole such that evaluation of one or the other alone undermined empirical efforts to support it. This had already been seen in the results of so-called "culture-fair" and "culture-free" tests of intelligence which were intentionally designed to minimize the effects of prior schooling or the need for previous learning but showed only modest correlations to more comprehensive tests and lower *g*-saturation on the whole (Cattell 1943, 1963; Horn 1965). If Gc must be measured given its centrality to the concept of intelligence, a more precise delineation of experiential differences in exposure and prior learning, above and beyond age and education, will be crucial to the generation of fair and equitable standards.

Horn's (1965) analyses presented evidence in support of a general visualization (Gv) primary factor that he described as "the processes of imagining the way objects may change as they move in space, maintaining orientation with respect to objects in space, keeping configurations in mind, finding the Gestalt among disparate parts in a visual field and maintaining a flexibility concerning other possible structurings of elements in space" (p. 310). Cattell had alluded to elements of so-called visual processing in his prior analyses but had mostly seen the issue as one related to the covariance observed in age-related decline of visual perceptual tasks due to degradation of visual acuity. As a developmental psychologist, Cattell understood and recognized the role that diminished vision played in attenuating test performances that relied on visual-perceptual processes, and thus he was inclined not to see it as an independent and distinct factor in its own right. Horn had no such difficulty and in his own elaborate style provided a strong argument for Gv including "quite definite support for the hypothesis of a general visualization dimension spanning the facets of Vz [visualization], S [spatial orientation], Cf [flexibility closure], Cs [speed closure], and DFT [adaptive flexibility] and dipping into measures of Gf when these involve Figural content" (pp. 280–281).

Horn (1965) presented similarly persuasive evidence in support of general speediness (Gs) as a primary factor. He hypothesized that Gs "could perhaps be an attribute indicating a state of test-taking effortfulness, rather than a stable trait" and subject to variation as a function of the nature of a given task (p. 310). Nonetheless, he indicated that Gs "is measured most purely in simple writing and checking tasks which require little in the way of complex relation-perceiving" and further admonished that "the function itself produces variance in the measure of most intellectual functions unless care is taken to cancel it out by measuring with both unspeeded and speeded tasks" (p. 310). Elements of this speed factor included Sc (speed copying), Wf (writing flexibility), and P (perceptual speed).

In a manner of speaking, Horn (1965) simply opened the floodgates with respect to the identification of primary factors, and as Cattell joined him in publication, the affirmation of the empirical support for an early Cattell-Horn Gf-Gc model was solidified (Horn and Cattell 1966a, b). As noted previously, it is a testament to Cattell's ability to alter his beliefs in the face of compelling evidence as it is clear that he was more than happy to let go once and for all the notion that there existed a general intelligence, whether comprised of a single *g* or two *g*'s (i.e., *g*f-*g*c). His outright rejection of this theoretical notion is given special credence by its conspicuous and prominent placement as the very first sentence ever published jointly between he and Horn (Horn and Cattell 1966a) which stated rather emphatically, "the theory of fluid and crystallized intelligence … seriously questions the notion that there is a unitary structure which can be designated general intelligence" (p. 253). Not content with upsetting the proverbial apple cart with this single premise, they add that Gf-Gc theory also "questions the belief—often implicit, but expressed clearly in a recent article by McNemar (1964)—that the conglomerate measured by combining subscores from a collection of intellectual tests is the best estimate of intelligence" (p. 253). In this regard, Horn and Cattell had become more aligned with Thurstone's (1946) original position that:

> Instead of attempting to describe each individual's mental endowment by a single index such as a mental age or an intelligence quotient, it is preferable to describe him in terms of a profile of all the primary factors which are known to be significant…. If anyone insists on having a single index such as an I.Q., it can be obtained by taking an average of all the known abilities. But such an index tends so to blur the description of each man that his mental assets and limitations are buried in the single index. (p. 110)[2]

[2] As yet another example regarding the influence of academic genealogy, Horn co-mentored this chapter's author's own dissertation which examined various methods of data aggregation in defining latent variables—that is, what is the best way to put two or more scores together mathematically to represent a single psychological construct. Prior to the final results, Horn predicted that a simple arithmetic average would emerge as the superior method, and in contrast to the manner in which many constructs and test score composites/clusters are calculated in the present day, he was indeed correct every bit as much as was Thurstone.

Neither Horn nor Cattell was alone in the rejection of a general factor. But for Cattell, whose life, education, and experiences were rooted in just such a notion, it must have been a truly radical departure from his early foundations. On the other hand, Horn was well acquainted with the "British" school and its ideology, having learned it firsthand from both Cattell and via his Fulbright work in Australia (1956–1957) and his position as a research associate at the University of London in 1972. Indeed, having "seen" the other side may have allowed Horn to view the British perspective from a more objective angle and therefore sensitized him to its shortcomings and allowed him to see problems and anomalies where others saw none. His insight into the problems with a general factor was revealed in an interesting manner more recently when he responded to a seemingly innocuous post on the first list serve ever created specifically dedicated to the topic of CHC theory by Kevin McGrew. Not long after its launch, a question arose regarding why the WJ-R (Woodcock and Johnson 1989) seemed to correlate so poorly with the venerable Wechsler scales and whether such modest correlations were indicative of a test that did not actually measure intelligence, particularly in the broad or general sense. Being a charter member of the list serve, Horn took the opportunity to address the issue—which at the time must have seemed that we had all returned to a state of quiet acquiescence of Spearman's *g*. Because these are his words (previously unpublished in their entirety), they merit inclusion in this chapter at this very point and are offered here for the benefit of the reader. His remarks, as they appear below, are quoted in their entirety, exactly as he wrote them on August 2, 1999.

> The problem is that there is no g--.e., no single g. Of course this is contrary to existing dogma. But dogma is dogma, not evidence, not something we want much of in science. It is an assumption implicitly accepted, an assertion made so frequently, by so many who are assumed to be (and assume themselves to be) authorities, and made so uncritically that it is widely accepted as true. (How could something said so often, so confidently, so casually, by so many, and so many smart, and informed people, not be true?) But the evidence adds up, as I have said now so many, many times –ad nauseam some may think. Still, for those who care about evidence, there is lots of it. One can examine it, and when one does one finds a drip, drip, drip of results from study after study punching out huge holes in the belief that there is a g (somewhere) and demonstrating that g hasn't been found and that it now seems unlikely that it will be found. In any case, if there is a g, we have yet to find it.
>
> And there is no contrary evidence, no evidence supportive of g. The only thing that gets treated as evidence is positive manifold of the intercorrelations among measures of cognitive abilities and a string of correlations with other variables that reflect this positive manifold. But this is evidence that Thurstone showed many years ago does not support a structural hypothesis of g, much less a developmental, genetic, neurological, educational, social, anthropological--n general, a construct validit--hypothesis.
>
> Recently, for example, McArdle has presented no fewer than three studies showing that g does not work structurally, developmentally, and dynamically. Also recently is the evidence of Richard Roberts analyses of the Armed Services Vocational Aptitude Battery (ASVAB)--he battery used in the data analysis parts of the infamous Herrnstein & Murray "Bell Curve" study. In the H&M studies obeisance was given to g when in fact the evidence of H&M's own factoring indicated no g. Roberts' results elegantly demonstrate this. (Here one might want to look at my review of H&M's book, there also pointing out this problem. Also good reading on this point are the Haut et al. reports, book and papers).
>
> Just a couple of years back, Schonemann and a whole host of responders to his work, concluded that there is no g. A little further back in history is Carroll's monumental work where, as I point out in several papers (again ad nauseum perhaps), there are no fewer than 8 different general factors, all quite distinct, but still referred to as "the" (singular) general factor or g. Prior to that, reviewing Jensen's "Bias in Mental Testing," Horn & Goldsmith found that what Jensen referred to as "g" in one chapter of his book was most similar to Gc, what he referred to as g in another chapter was similar to Gf and what he accepted as g in still another chapter was essentially Gv. While Jensen's work presents particularly stark examples of this chameleon-like interpretation of ability measurements, in fact he is simply doing what many others do. But if one looks at the evidence, s/he will see Gf, Gc, Gv, etc., have quite distinct construct validities –quite different relationships to neurological, educational, vocational, genetic variables--n general the network of variables that provide a basis for understanding human capabilities. Going back further yet there are the classic studies of El Koussy (1935) and Rimoldi (1948) studies that

were steadfastly and beautifully designed to prove the validity of g, but concluding--eluctantly, almost sadl--that the g hypothesis can not be supported. Then, too, more tangentially, there are the well-designed and well-executed studies of Gustafsson (1984, 85), Undheim (1976) and Undheim and Gustafsson (1987) showing that in batteries of tests designed for children the general factor was identical to the Gf factor. Relatedly, there are the results from our work showing that a battery very carefully designed to provide evidence of one and only one factor corresponding to Spearman's g comes very close to succeeding, as in Thurstone and Thurstone (1941), but the g that is indicated is Gf (and only Gf): as soon as other well-regarded indicators of intelligence, such as those of Gc, Gv, Ga, etc., are considered, that "g" factor (which is Gf) disappears.

So, old friends of NASP--and Jensen, Bouchard, Eysenck, Carroll, etc. (the list is long) be aware that no battery of ability tests provides a measure of g, because there is no g, only conglomerate, composites. One good reason why the composite score on the WJR may correlate at a lower level with this or that composite or other variable is because the WJR is well designed to provide measures of the different concepts of human cognitive capabilities--hat in the vernacular is referred to as intelligenc--that, so far, have been indicated by research. It provides measures of Gv, Ga, Gs, SAR, TSR, Gq, as well as Gf and Gc, whereas the Wechsler scales provide only Gc, a Gf-Gv mixture, and a very weak SAR and Gs. The composite of the WJR is thus a broad mixture relative to the composites of other, more narrow, batteries. But even when the same elements of a composite appears in other batteries, the elements enter in different proportions to the whole: the composites of different published tests differ not only in breadth, but also in the proportions of different abilities that contribute to the composite measure. A broad composite relative to a narrow one needn't necessarily correlate at a lower (or higher) level with other variables, but it may. It depends on breadth of the variable with which the composites are correlated and on how well the components of a predictor composite match the components of the predicted variable. In prediction of job performance on many jobs, for example, as in much of the research reviewed by Schmidt and Hunter (1992), the broad composite of the ASVAB (and it's descendents) predicts better than most narrower composites.

Enough said. These points are not highly debatable. One can have a quibble here and there, but basically the evidence at this point in history doesn't permit much deviation from the conclusion that there is no g; the emperor is naked.

Sincerely,
John Horn

Of primary significance in Horn's comments is the argument that attempts to distill measurement of a single general ability factor result only in the measurement of a distinct broad ability factor, notably Gf, when language and cultural elements are not used, or Gc, when they are. Horn also makes reference to the version of the theory that became known as the "modern Gf-Gc" framework and which greatly expanded the number of primary (narrow) and secondary (or broad, but not general) mental abilities included in the theory. That expansion was relatively rapid paced as in (1966a), only a year after introducing two new broad, secondary factors in his dissertation (Gv and Gs), he adds a fifth to the collection "F" (general fluency) which was more of a process that was reflected in certain tasks, particularly those that required speed. The specific abilities subsumed under this broad factor included Fa (associated fluency), Fi (ideational fluency), and Fw (word fluency). By the early 1990s, Horn had expanded the broad abilities which now numbered ten in all and included as many as 80 or so primary (narrow) abilities subsumed by them collectively. The usual suspects remained, Gf, Gc, Gv, and Gs, but F had become TSR (fluency of retrieval from long-term storage), and another memory component had been identified, SAR (short-term apprehension and retrieval). In addition, three new broad abilities were incorporated, one very similar to Gv but relying on auditory stimuli and perception, Ga (auditory processing); one very similar to Gf and Gc, but specific to mathematical or quantitative knowledge (Gq); and one very similar to speed (Gs) but labeled correct decision speed (CDS) to indicate rapidity in providing correct, not merely quick, responses in relatively simple comprehension, reasoning, or problem-solving tasks. A final factor related to reading and writing skills, Grw, was also added as it became more commonplace to include published tests of academic achievement right alongside cognitive ability tests (a distinction that Horn indicated was only a semantic issue and not a true difference). As Horn continued to advance the theory, it began to be referred to as "modern" Gf-Gc theory and later as "extended" Gf-Gc theory.

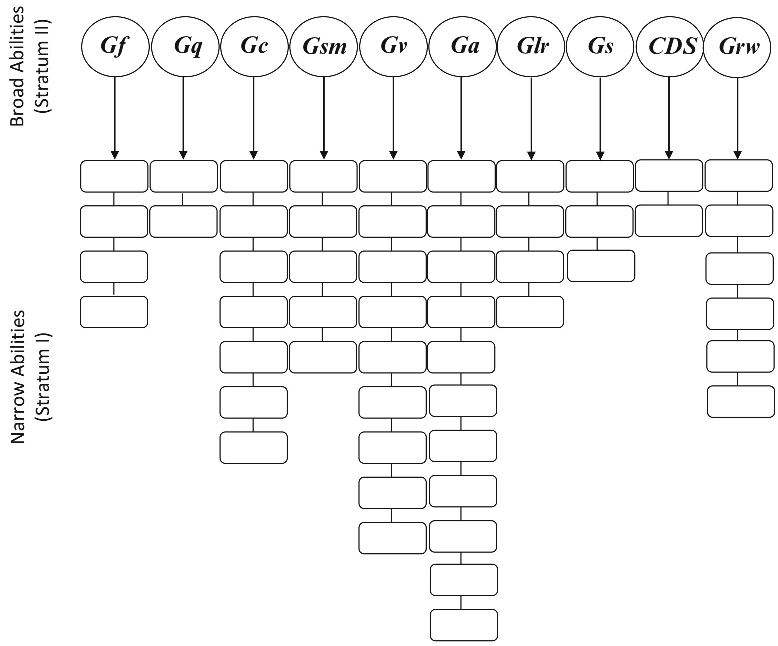


Fig. 15.3 The Cattell-Horn "extended" Gf-Gc model

Regardless of the name, a general factor supreme to all other broad and narrow abilities was, of course, absent in each refinement of the Gf-Gc model. Figure 15.3 provides an illustration of the extended Gf-Gc model delineated by Horn late in his career (Horn and Blankson 2007; Horn and Noll 1997).

Around 1988, having a rather a theoretical test of intelligence under his belt already (i.e., Woodcock-Johnson Psycho-Educational Battery; Woodcock and Johnson 1977), Richard Woodcock journeyed to the University of Southern California likely in search of better theoretical guidance. His collaboration with Horn resulted in not only the first published test specifically designed to operationalize modern Gf-Gc theory, it also prompted the publication of an often overlooked and important article that further established the validity and utility of the theory, particularly as a platform for test development (Woodcock 1990) that not only affirmed the empirical support for modern Gf-Gc theory across a variety of published instruments, but also reintroduced a concept that had long since vanished—that "the clinician may find it helpful to 'cross' batteries to obtain a set of measurements required for a particular assessment" (p. 252). Coupled with his three prior points on appropriate psychometric principles that must govern test data (i.e., knowledge of the factorial composition of each subtest in a battery, formation of broad ability composite or cluster scores drawn from two similar but qualitatively distinct narrow ability indicators, and avoidance of subtests with mixed-factor loading that complicate, if not obviate, explanations of test performance), the foundations of what would become CHC cross-battery assessment (McGrew and Flanagan 1998; Flanagan and Ortiz 2001; Flanagan et al. 2007, 2013) are evident. Ironically, the suggestion that batteries could be "crossed" was not a new idea. Cattell (1943) himself, commenting on the vast number of published intelligence batteries generated over the preceding 20 years, noted that "it is at first glance a sufficiently impressive window display; indeed, the adult-testing psychometrist may be led to consider himself richly equipped with 44 tests" (p. 154). Using a variety of tests culled from a range of individual batteries was relatively commonplace and even necessary precisely because they lacked any real theoretical foundation. According to Atwell et al. (1941), "the verbal items of the Wechsler-Bellevue Intelligence Scale, supplemented by a vocabulary test and alternate arithmetic questions from the Alpha test would be satisfactory … it was expected that the Beta Block Counting test, the 11, 12, and 14 year levels of the Porteus mazes, and the third, fourth, and sixth designs of the Wechsler-Bellevue Block Design test would be relatively useful" (p. 898). Both Cattell's and Horn's analyses in support of the theory of fluid and crystallized intelligence relied on test data generated across a range of tests and batteries. For unknown reasons, however, clinical practice with testing had settled largely into a test-kit-driven modality where, apart from some of the practices of neuropsychologists, intelligence was measured and determined almost exclusively by the structure of a single test—a practice fraught with peril, as Cattell, Horn, and Woodcock all noted, because it meant that the resulting intelligence was not equivalent to the prevailing conceptualization of intelligence but rather "an average of whatever has been chosen by the test author to be included in that battery" (p. 250). Whereas the field in general might have lost sight of this principle, cautions remained regarding the inappropriate use of IQ and inaccurate perceptions of the construct as eloquently stated by Salvia and Ysseldyke, "different intelligence tests are simply samples of behaviors. For that reason it is wrong to speak of a person's IQ. Instead, we can refer only to a person's IQ on a specific test … Because the behavior samples are different for different tests, one must always ask, 'IQ on what test?.'" (p. 158). Such wisdom has gone unheeded too long in applied psychology, but the popularity and rise of CHC theory and the relatively quick mustering by test developers in aligning themselves with current intelligence theory have reawakened an interest in understanding this important idea while at the same time helping to de-emphasize the utility and value of global intellectual scores.

CHC Theory and Carroll: Coming of Age

Adolescence is often a period of time where concerns about self-identity become salient. And as the Cattell-Horn modern Gf-Gc theoretical framework matured, an important event occurred that provided an additional foundation regarding the evolution of CHC theory into a form more closely resembling its present incarnation. That event was the publication of John Carroll's (1993) remarkable book *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. In short, Carroll undertook examination of over 50 years worth of data drawn from tests of intelligence and cognitive abilities to create a single, massive database with which he could perform his own factor analyses. Carroll was insistent upon the use of exploratory rather than confirmatory techniques because he wished to allow the "data to speak for themselves." Whether he went in with any preconceptions about the structure of intelligence or not is unknown, although his eventual insistence and firm arguments in favor of a general factor may offer some insight into his a priori beliefs. Nevertheless, Carroll distilled some 2,000 available datasets published between 1983 and 1985 (approximately) down to about 460 which he believed met critical criteria necessary to permit his analyses and required for the sake of eliminating data that were less than trustworthy on various psychometric and theoretical grounds. This included datasets that contained a substantial number of variables, a sizable sample (greater than 100), a published correlation or covariance matrix, and sufficient descriptions of the sample and variables to permit interpretation (Carroll 1997).

Carroll drew from Cattell (1971) in his use of the term "stratum" to describe the difference between first-, second-, and third-order factors. Because different datasets may contain information that only permits construction of only a first- and second-order factor, there had been some confusion about whether a general factor was indicated at the second level or at the third.

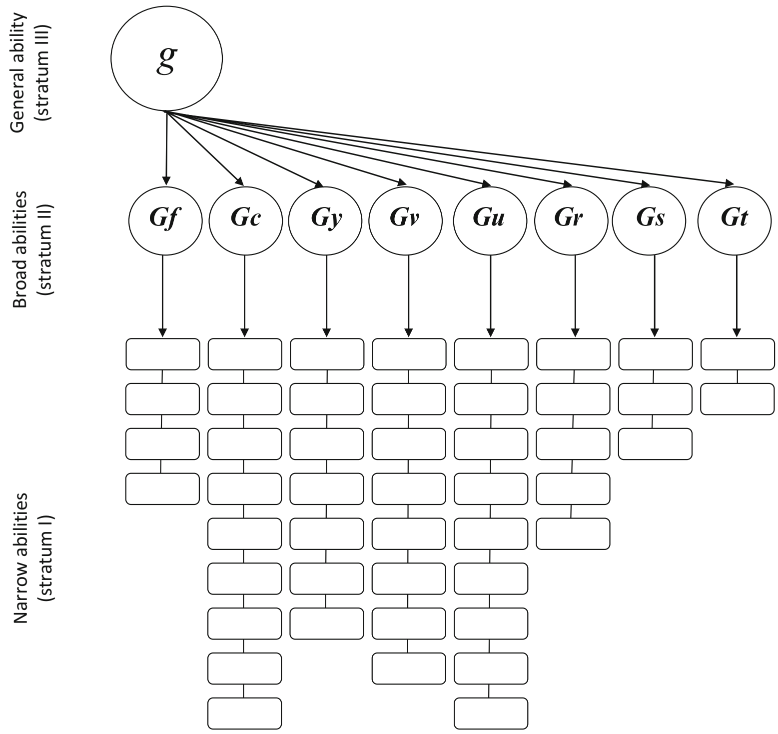


Fig. 15.4 Carroll's three-stratum model

Carroll sought to minimize such confusion by implementing Cattell's stratum notion in ways that clearly delineated the nature of the ability or factor identified at each of the three levels. Thus, Carroll (1997) named the "first-order factors resulting from analysis of typical sets of psychological datasets *factors at the first stratum,* or *stratum I factors*" (p. 124; emphasis in original). Likewise, "*stratum II factors* were second-order factors from such data sets, and *stratum III factors* were third-order factors from such datasets" (p. 124) and postulated that any ability must fall at one of these three levels.

Carroll's (1997) ultimate intent was to provide extensions to and expansions of existing theoretical models by providing a "provisional statement about the enumeration, identification, and structuring of the total range of cognitive abilities known or discovered thus far" (p. 124). If taken at face value, Carroll's work provides significant support for modern Gf-Gc theory, above and beyond other conceptualizations of intelligence. Carroll (1993) noted specifically that "the Cattell-Horn model … is a true hierarchical model covering all major domains of intellectual functioning … among available models it appears to offer the most well-founded and reasonable approach to an acceptable theory of the structure of cognitive abilities" (p. 62). Despite this ringing endorsement, Carroll outlined specific areas in which his proposed three-stratum model differed from the Cattell-Horn formulation and other similar theoretical models. In doing so, it may have fostered the impression that he was indeed, despite his reservations, proposing a newer and better model which should, in logical fashion, supplant the less correct extended Gf-Gc framework. Whatever the reason, Carroll's formulation has become known in its own right as the three-stratum model of intelligence. Figure 15.4 provides an illustration of the model that Carroll proposed on the basis of his analyses.

At the narrow ability level, Carroll's (1993) analyses revealed about 65 primary, first-order factors. These are specific tasks measured relatively purely and cleanly by a single task, most often couched as a subtest within a larger battery of tests. The sheer number of abilities at this level precludes inclusion of their names and definitions in their entirety, and the reader is referred to other sources for such information (e.g., Carroll 1997). At the broad ability level, Carroll's (1993) analyses revealed only eight broad abilities—a result that was not surprising since he likely confined himself to intelligence tests. If the nomenclature applied by Carroll is loosely interpreted, the resemblance to modern (or extended) Gf-Gc theory was quite good. For example, the broad abilities of Gf, Gc, Gv, and Gs were equivalent in both definition and labels. The Cattell-Horn model used Ga for auditory processing, whereas Carroll used the designation "Gu." A greater difference was evident in Carroll's use of "Gy" for general memory and learning, in contrast to the Horn-Cattell use of SAR (which began being referred to as Gsm), as well as in the delineation of "Gr" or broad retrieval ability which was called TSR (which also began being referred to as Glr) in the extended Cattell-Horn Gf-Gc model. And finally, Carroll labeled his speed factor "Gt" referring to processing speed and RT (decision speed) as one unified broad factor compared to the Cattell-Horn formulation that split the broad factor into two components: Gs, processing speed, and CDS, correct decision speed.

The most significant and obvious difference in Carroll's three-stratum model and the extended Horn-Cattell Gf-Gc model rests with the concept of a general factor. As already discussed, Cattell had envisioned a general factor comprised of two distinct components, and Horn had dispensed with the notion of a general factor from the outset. Neither saw any rationale nor need for it, and neither factored their data in a way that supported it. For them, it was a mathematical nonnecessity, a statistical artifact, and an illusion that simply was not worth pursuing. Carroll, on the other hand, appeared to care deeply about a general factor. In adopting a three-stratum model vs. a two-stratum model, he made a deliberate choice to create a level in which the general factor could exist. In his diagrams, this third-order factor (which should have less significance than first- or second-order factors) is also intentionally placed at the top providing a clear indication of its superiority within a hierarchical structure. Moreover, it is positioned to the left side right above Gf and Gc, as a way of indicating that those abilities bear the closest relationship to it by virtue of being the most *g*-saturated. Of course, one could argue, as did Cattell, Horn, Thurstone, and others, that the general factors are simply Gf anyway and a mere duplication of the second-order factor. In 2003, he asserted that "researchers who are concerned with this structure in one way or another…. can be assured that a general factor *g* exists, along with a series of second-order factors that measure broad special abilities" (p. 19). It seems impossible, perhaps even needless, to elaborate more on this debate. It is sufficient to recognize that it is in fact a debate, and not to assume as so many have done erroneously, that *g* is an established scientific fact. Likewise, if one wishes to dispense with *g* altogether, it should only be in service to the purposes of evaluation which neither need nor require it. The peril in this regard is to succumb wholly to one position or the other when the reality is that neither, as of yet, is supported by any incontrovertible evidence that makes either position defensible. If applied psychologists continue to demand *g*, or global, full-scale scores intended to represent general intelligence, then test publishers will no doubt continue to offer it. Whatever the case, it should be recognized that aggregating scores to produce a general index is a simple enterprise, perhaps too simple, and that because it can be so easily constructed, it may prove seductive for use in ways that cannot be supported psychometrically and theoretically.[3]

[3] I once had the opportunity to ask Richard Woodcock why his WJ-R contained a global intellectual ability (GIA) score when the test was operationalized according to extended Gf-Gc theory that did not specify such a general factor. His response was both enlightening and discouraging as he admitted, almost apologetically and reluctantly that he could not have gotten the test published without including it.

CHC Theory and an Integrated Framework: Maturing in Adulthood

Despite the apparent differences between the extended Cattell-Horn Gf-Gc model and Carroll's three-stratum model, as well as the irreconcilable issue of *g* or not to *g*, a rather unusual thing happened in the late 1990s. Whereas it would have been logical to reason that intelligence theorists and their personal frameworks would fracture the landscape to the point that none of them would be suitable for guiding test development, the very opposite occurred, and attempts were made to consolidate and integrate them in one overarching theoretical framework. Why this may have occurred can probably be traced to several factors. For example, neither extended Gf-Gc nor the three-stratum model was well recognized or popular in the mainstream discussions or measurement of intelligence, so not a lot of researchers, let alone practitioners, paid them much attention. In addition, the WJ-R was making strong inroads, however, and beginning to craft out a place of its own among the now relatively small pantheon of comprehensive intelligence tests that had boiled down, more or less, to the Wechsler and Stanford-Binet scales, Kaufman Assessment Battery for Children, and Differential Abilities Scales. Yet it still lagged significantly behind the others, particularly the Wechsler scales in use and acceptance among applied psychologists. As noted previously, Kevin McGrew and Dawn Flanagan adopted the principles set forth by Woodcock (1990) and outlined the foundations of cross-battery assessment (often referred to as XBA; Flanagan and McGrew 1997; McGrew and Flanagan 1998) which placed extended Gf-Gc theory at its core. But perhaps the key circumstance during this period was McGrew's reconciliation and integration of the Cattell-Horn Gf-Gc and Carroll's three-stratum model as a guide to bridging intelligence theory with the practice of intellectual assessment (McGrew 1997). McGrew specifically outlined a "proposed comprehensive Gf-Gc framework" that he generated via joint confirmatory cross-battery factor analyses drawn from subtest data generated by the DAS, K-ABC, KAIT, SBIV, WISC-III, WPPSI-II, WAIS-III, and WJ-R.

Given the similarities of the two models, it would seem that such a task did not present many difficulties, but that would be an incorrect characterization. First, many of the subtests and batteries McGrew included had never been factored in accordance with Gf-Gc theory, as had the WJ-R subtests. Due to the confirmatory nature of the analyses, expert classifications had to be made for many tests to provide a foundation for testing the relations among them. Second, not all batteries measured all broad abilities, and thus, data on some abilities was likely to be rather limited (e.g., only the WJ-R provided measures of auditory processing). But as Cattell had noted, the theory should be tested in its totality, and it would be necessary to ensure that adequate coverage in terms of both breadth of abilities measured as well as depth of abilities existed or else the results might prove rather limited.

McGrew's initial analysis resulted in the extraction of nine broad abilities (one less than Cattell-Horn's and one more than Carroll's) including Gc, Gv, Gq, Grw, Gsm, Gf, Gs, Glr, and Ga. Given the limitations of the tests he used, he was only able to identify about 36 narrow abilities in the dataset. Some, like Gsm, Gf, and Gs, contained only two identified narrow abilities, whereas others like Gc, Gv, and Grw subsumed numerous narrow abilities, seven, six, and seven, respectively. Of particular note is that in many cases (9 out of the 36), only one subtest from one battery was present to provide data for a given narrow ability. This is not so much a limitation of McGrew's study but more an indictment on the narrowness of the factor structures upon which the tests had been developed. Later versions of the model contain ten broad abilities that include Gt (reaction time/decision speed) which subsumes three additional narrow abilities, but it is sometimes omitted since it is not currently measured by any major intelligence battery. In addition, some researchers have called for a separation of Grw into its constituent components, using the designations Grw-R for reading skills and Grw-W for writing skills (Flanagan and Ortiz 2001; Flanagan et al. 2007, 2013). McGrew asserts that these factors do not distinguish themselves in factor analytic studies and tend to be very highly correlated. The association between the two is quite clear as they are both symbolic aspects of language development. But because schooling specifically treats them and teaches them as independent skills, there is a practical advantage in assessment that comes from making this distinction. However, it has not yet been fully incorporated into current CHC theoretical specifications as it presents a condition that cannot be supported empirically.

Once again, it appears that fortuitous events reopened the floodgates regarding development and advancement of the theory. In hindsight, the integrated model was only just beginning to blossom. Consider that in 1985, Horn, Carroll, and McGrew were called upon as consultants regarding the revision of the 1978. Schneider and McGrew (2012) hails this "meeting of the minds" as "a moment where the interests and wisdom of a leading applied test developer (Woodcock), the leading proponent of Cattell-Horn Gf-Gc theory (Horn), and one of the preeminent educational psychologists and scholars of the factor analysis of human abilities (Carroll) intersected" came into a perfect alignment of psychometric stars that served, in his opinion, as "the flash point that resulted in all subsequent theory-to-practice bridging events leading to today's CHC theory and related assessment developments" (p. 144). Following his later integration, McGrew teamed with Woodcock and Nancy Mather in 1999 to develop and subsequently publish the Woodcock-Johnson III (Woodcock et al. 2001). By this time, the integrated model had taken on a new name that McGrew cites as having first been presented in the WJ-R III technical manual (McGrew and Woodcock 2001) where the full name "Cattell-Horn-Carroll theory of cognitive abilities" appears and is described as "an amalgamation of two similar theories about the content and structure of human cognitive abilities" (p. 9). McGrew further notes that by amalgamation, and by implication, the name was partially derived from personal communication with both Horn and Carroll in 1999 and that they had recommended the name and that it be structured according to order of contribution.

And so was born CHC theory, which by this point continued to provide the long sought-after common taxonomy of abilities and nomenclature that permitted further advancement by ensuring that all researchers and all studies were finally talking about the same things. What was Gf in one study or on one test was now comparable and relatively equivalent to Gf in another study or another test. It is no small point to say that the viability and success and ever growing popularity of CHC theory is rooted largely in the fact that it has created a landscape that has enhanced theoretical and psychometric development by clearly delineating factors, tests that measure those factors, tests that do not measure them, and language common to all that reduces confusion and ambiguity. McGrew referred to Carroll's work, upon which much of the nomenclature of the integrated CHC model is based, as the "Rosetta stone" of the field of intelligence. With an extensive history of development and increasingly widespread use in current research and development, the momentum of CHC theory is such that it may sweep up many other lesser models in its wake. This fact has not gone unnoticed by test publishers either in that the first decade of the twenty-first century saw a rapid convergence on CHC theory or radical revisions to the factorial structure of tests and their subtests to the point that it was described, albeit facetiously, as a miraculous and "sudden collective psychometric epiphany" (Flanagan et al. 2007, p. 221).

Without question, CHC theory has had a significant impact in terms of promoting theoretical developments in the area of intelligence and cognitive abilities going on nearly a full century now. And despite getting an extremely late start with respect to influencing the design and structure of tests used in measuring human cognitive abilities, it has made impressive gains in just the last two decades and shows little sign of slowing down any time soon. Figure 15.5 provides an

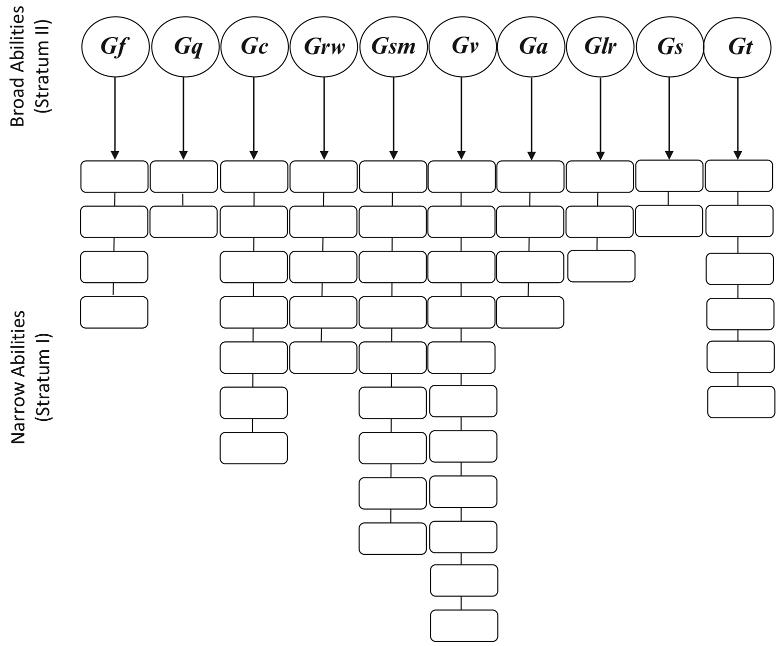


Fig. 15.5 CHC theory "early" integrated model

illustration of the "early" integrated CHC model as formulated by McGrew (1997).

CHC Theory—Aging Gracefully, A Summary of Sorts

With incredible modesty and recognition of the absolute purity of the scientific process, note with all due sincerity that "the extended theory of fluid and crystallized (Gf and Gc) cognitive abilities is wrong, of course, even though it may be the best account we current have of the organization and development of abilities thought to be indicative of human intelligence" (p. 41). They further assert that "all scientific theory is wrong. It is the job of science to improve theory. That requires identifying what is wrong with it and find out how to change it to make it more nearly correct" (p. 41). It is unlikely that the CHC theory will ever be modified to the point that it becomes completely correct, and it may well be overthrown by a better and more accurate theory at some point in the future. CHC theory offers just one view of the nature and specifics of human cognitive abilities, including intelligence, but it need not be seen as the final word on the matter. Nevertheless, its value as a useful and guiding paradigm is well established and continues to inspire both confidence and interest in what might be accomplished with further refinement. Consider that in the very latest incarnation of CHC theory (Schneider and McGrew 2012), fully 16 broad abilities are now specified along with 81 narrow abilities. In addition to the ten abilities described previously, Schneider and McGrew outline six new ones including Gkn (domain-specific knowledge), Go (olfactory ability), Gh (tactile ability), Gp (psychomotor ability), Gk (kinesthetic ability), and Gps (psychomotor speed). Despite this elegant expansion, none of the aforementioned broad abilities are measured by any current battery which precludes their measurement and

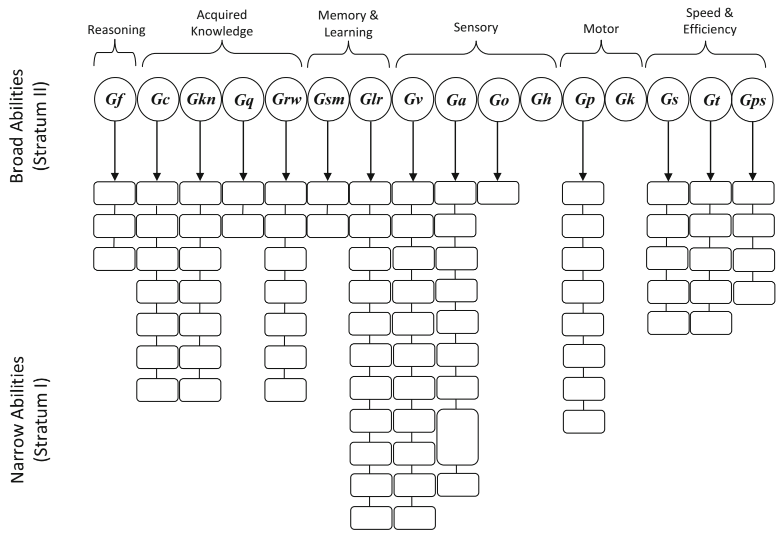


Fig. 15.6 CHC theory "current" integrated model

limits assessment to the ten abilities noted previously. Likewise, of the 81 narrow abilities, only 38 (just two more than before) can be measured and evaluated by current tests. A simplified illustration of this "current" version of the integrated CHC model is presented in Fig. 15.6.

Interestingly, much of the current discussion regarding CHC theory does so without much attention to the notion of general intelligence. There does not appear to be any deliberate attempt to include or exclude *g* in the latest integrated CHC theoretical models, and its relative absence may be more of a reflection regarding interest in broad and narrow abilities and their ability to predict well specific areas of academic skill development above and beyond what is captured solely by *g* (Keith and Reynolds 2010). Because the preceding discussion has made it clear that *g* is neither necessary nor desirable in all cases, and because no test publisher is likely to release a test without including some global score, it will remain an option for any assessor who wishes to use it. Therefore, its inclusion within current theory is not as compelling as it once might have been, especially as research into the narrow abilities continues to demonstrate important and specific relationships between cognitive and academic abilities that are quite useful in psychoeducational evaluations (Schneider and McGrew 2012). While the theory-practice gap that Cattell, Horn, and McGrew intended to bridge via refinements to their theoretical models remains even today, it is now a combination of tests that have not yet caught up to theory coupled with theory that continues to grow and refuses to sit still. In that regard, Cattell, Horn, and Carroll would all likely take heart in knowing that their burning interest in furthering an understanding of intelligence and human cognitive abilities has been passed on intact to the current generation of psychological scientists. CHC theory has grown up, and it would be safe to assume that Cattell, Horn, and Carroll would likely be very proud to know that their goals, their ideas, and their passion are still alive and kicking in CHC theory to the present day.

References

Atwell, C. R., Bloomberg, W., & Wells, F. L. (1941). Psychometrics at an army induction center. *New England Journal of Medicine, 224*, 898–899.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Carroll, J. B. (1997). Commentary on Keith and Witta's hierarchical and cross-age confirmatory factor analysis of the WISC-III. *School Psychology Quarterly, 12*, 108–109.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38*, 592.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 2*, 153–193.

Cattell, R. B. (1950). *Personality: A systematic theoretical and factual study*. New York: McGraw-Hill.

Cattell, R. B. (1957a). *Personality and motivation structure and measurement*. New York: Harcourt, Brace & World.

Cattell, R. B. (1957b). *The IPAT culture fair intelligence scales*. Champaign: Institute for Personality and Ability Testing.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1–22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

El Koussy, A. A. H. (1935). The visual perception of space. *British Journal of Psychology,* Monograph Supplement, *20*, 1–80.

Flanagan, D. P., & McGrew, K. S. (1997). A cross-battery approach to assessing and interpreting cognitive abilities: Narrowing the gap between practice and science. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 314–325). New York: Guilford Press.

Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. New York: Wiley.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd ed.). New York: Wiley.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd ed.). New York: Wiley.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Psychological Bulletin, 61*, 401–404.

Horn, J. L. (1965). *Fluid and crystallized intelligence: A factor analytic and developmental study of the structure among primary mental abilities*. Unpublished doctoral dissertation, University of Illinois, Champaign.

Horn, J. L., & Blankson, N. (2007). Foundations for a better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (2nd ed., pp. 41–68). New York: Guilford Press.

Horn, J. L., & Cattell, R. B. (1966a). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107–129.

Horn, J. L., & Cattell, R. B. (1966b). Age differences in primary mental abilities. *Journal of Gerontology, 21*, 210–220.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 53–91). New York: Guilford Press.

Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools, 47*(7), 635–650.

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 151–179). New York: Guilford Press.

McGrew, K. S., & Flanagan, D. P. (1998). *The Intelligence Test Desk Reference (ITDR): Gf-Gc cross-battery assessment*. Boston: Allyn & Bacon.

McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock-Johnson III technical manual*. Itasca: Riverside.

McNemar, Q. (1964). Lost: Our intelligences. Why? *American Psychologist, 19*, 871–872.

Rimoldi, H. J. (1948). Study of some factors related to intelligence. *Psychometrika, 13*, 27–46.

Schmidt, F. L., & Hunter, J. E. (1992). Development of a causal model of processes determining job performance. *Current Directions in Psychological Science, 1*, 89–92.

Schneider, J., & McGrew, K. S. (2012). The Cattell-Horn-caroll model of Intelligence In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (3rd ed.). (pp. 99–143) New York: Guilford.

Thurstone, L. L. (1946, February). Theories of intelligence. *Scientific Monthly* (pp. 101–112).

Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence (Psychometric Monographs, No. 2)*. Chicago: University of Chicago Press.

Undheim, J. O. (1976). Ability structure in 10-11-year-old children and the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 68*, 411–423.

Undheim, J. O., & Gustafsson, J. E. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research, 22*, 149–171.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.

Wells, F. L. (1932). Army alpha revised. *Journal of Personality, 10*, 411–417.

Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment, 8*, 231–258.

Woodcock, R. W., & Johnson, M. B. (1977). *The Woodcock-Johnson Psycho-Educational Battery*. Chicago: Riverside Publishing.

Woodcock, R. W., & Johnson, M. B. (1989). *The Woodcock-Johnson Psycho-Educational Battery—Revised*. Chicago: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca: Riverside Publishing.

Multiple Intelligences in the New Age of Thinking

16

Robert J. Sternberg

In this chapter, I describe two theories that involve the idea of "multiple intelligences." The first is the theory of multiple intelligence (MI theory—Gardner 1983, 1993, 2006). The second is the theory of successful intelligence (Sternberg 1997, 2003, 2005, 2009, 2010).

Multiple Intelligences Theory

Howard Gardner (1983, 1993, 2006) has proposed a theory of multiple intelligences, according to which intelligence comprises multiple independent constructs, not just a single, unitary construct. However, instead of speaking of multiple abilities that together constitute intelligence (e.g., Thurstone 1938), this theory distinguishes eight distinct intelligences that are relatively independent of each other.

The multiple intelligences are linguistic intelligence, used to read, write, speak, and listen to speech; logical-mathematical intelligence, used to solve mathematical and logical problems; spatial intelligence, used to imagine how objects would look if they were rotated or otherwise displaced in space; musical intelligence, used to compose, play, and appreciate music; bodily-kinesthetic intelligence, used to coordinate oneself and to participate successfully in athletics; interpersonal intelligence, used to understand other people; intrapersonal intelligence, used to understand oneself; and naturalist intelligence, used to understand the natural world.

Each intelligence is a separate system of functioning, although these systems can interact to produce what we see as intelligent performance. Looking at Gardner's list of intelligences, you might want to evaluate your own intelligences, perhaps rank ordering your strengths in each.

In some respects, Gardner's theory sounds like a factorial one. It specifies several abilities that are construed to reflect intelligence of some sort. However, Gardner views each ability as a separate intelligence, not just as a part of a single whole. Moreover, a crucial difference between Gardner's theory and factorial ones is in the sources of evidence Gardner used for identifying the eight intelligences. Gardner used converging operations, gathering evidence from multiple sources and types of data.

In particular, the theory uses eight "signs" as criteria for detecting the existence of a discrete kind of intelligence (Gardner 1983, pp. 63–67):

1. Potential isolation by brain damage. The destruction or sparing of a discrete area of the brain (e.g., areas linked to verbal aphasia) may destroy or spare a particular kind of intelligent behavior.

R.J. Sternberg (✉)
Department of Human Development, B44 MVR, Cornell University, Ithaca, NY 14850, USA
e-mail: Robert.sternberg@cornell.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_16,

2. The existence of exceptional individuals (e.g., musical or mathematical prodigies). They demonstrate extraordinary ability (or deficit) in a particular kind of intelligent behavior.
3. An identifiable core operation or set of operations (e.g., detection of relationships among musical tones). It is essential to performance of a particular kind of intelligent behavior.
4. A distinctive developmental history leading from novice to master. It is accompanied by disparate levels of expert performance (i.e., varying degrees of expressing this type of intelligence).
5. A distinctive evolutionary history. Increases in intelligence plausibly may be associated with enhanced adaptation to the environment.
6. Supportive evidence from cognitive-experimental research. An example would be task-specific performance differences across discrete kinds of intelligence (e.g., visuospatial tasks versus verbal tasks). They would need to be accompanied by cross-task performance similarities within discrete kinds of intelligence (e.g., mental rotation of visuospatial imagery and recall memory of visuospatial images).
7. Supportive evidence from psychometric tests indicating discrete intelligences (e.g., differing performance on tests of visuospatial abilities versus on tests of linguistic abilities).
8. Susceptibility to encoding in a symbol system (e.g., language, math, musical notation) or in a culturally devised arena (e.g., dance, athletics, theater, engineering, or surgery as culturally devised expressions of bodily-kinesthetic intelligence).

Gardner does not dismiss entirely the use of psychometric tests. But the base of evidence used by Gardner does not rely on the factor analysis of various psychometric tests alone. In thinking about your own intelligences, how fully integrated do you believe them to be? How much do you perceive each type of intelligence as depending on any of the others?

Gardner's view of the mind is modular. Modularity theorists believe that different abilities—such as Gardner's intelligences—can be isolated as emanating from distinct portions or modules of the brain. Thus, a major task of existing and future research on intelligence is to isolate the portions of the brain responsible for each of the intelligences. Gardner has speculated as to at least some of these locales. But hard evidence for the existence of these separate intelligences has yet to be produced. Furthermore, there is no real evidence for the strict modularity of Gardner's theory. Consider the phenomenon of preserved specific cognitive functioning in autistic savants. Savants are people with severe social and cognitive deficits but with corresponding high ability in a narrow domain. They suggest that such preservation fails as evidence for modular intelligences. The narrow long-term memory and specific aptitudes of savants may not really be intelligent. Thus, there may be reason to question the intelligence of inflexible modules.

I do not detail this theory further because there has been no empirical evidence collected since MI theory was proposed that validates the theory as a whole and the one extensive study that has been done yielded results inconsistent with it (Visser et al. 2006).

The Triarchic Theory of Successful Intelligence

The Nature of Intelligence

There are many definitions of intelligence, although intelligence is typically defined in terms of a person's ability to adapt to the environment and to learn from experience (Inteligence and its Measurement: A Symposium, 1921; Sternberg and Detterman 1986). The definition of intelligence here is somewhat more elaborate and is based on my (Sternberg 1984, 1997, 1998, 1999b, 2000, 2003) theory of successful intelligence. According to this definition, (successful) intelligence is (1) the ability to achieve one's goals in life, given one's sociocultural context; (2) by capitalizing on strengths and correcting or compensating for weaknesses; (3) in order to adapt to, shape, and select environments; (4) through a combination of analytical, creative, and practical abilities. In recent years, I have emphasized that intelligence best serves individuals and societies when it is augmented by wisdom (Sternberg 2003, 2008).

According to the proposed theory of human intelligence and its development (Sternberg 1997, 1999a, 2003, 2004, 2009), a common set of processes underlies all aspects of intelligence. These processes are hypothesized to be universal. For example, although the solutions to problems that are considered intelligent in one culture may be different from the solutions considered to be intelligent in another culture, the need to define problems and translate strategies to solve these problems exists in any culture. Even within cultures, there may be differences in what different groups mean by intelligence (Okagaki and Sternberg 1993; Sternberg 1985b).

Metacomponents, or executive processes, plan what to do, monitor things as they are being done, and evaluate things after they are done. Examples of metacomponents are recognizing the existence of a problem, defining the nature of the problem, deciding on a strategy for solving the problem, monitoring the solution of the problem, and evaluating the solution after the problem is solved.

Performance components execute the instructions of the metacomponents. For example, inference is used to decide how two stimuli are related and application is used to apply what one has inferred (Sternberg 1977). Other examples of performance components are comparison of stimuli, justification of a given response as adequate although not ideal, and actually making the response.

Knowledge-acquisition components are used to learn how to solve problems or simply to acquire declarative knowledge in the first place (Sternberg 1985a). Selective encoding is used to decide what information is relevant in the context of one's learning. Selective comparison is used to bring old information to bear on new problems. And selective combination is used to put together the selectively encoded and compared information into a single and sometimes insightful solution to a problem.

Although the same processes are used for all three aspects of intelligence universally, these processes are applied to different kinds of tasks and situations depending on whether a given problem requires analytical thinking, creative thinking, practical thinking, or a combination of these kinds of thinking. In particular, analytical thinking is invoked when components are applied to fairly familiar kinds of problems abstracted from everyday life. Creative thinking is invoked when the components are applied to relatively novel kinds of tasks or situations. Practical thinking is invoked when the components are applied to experience to adapt to, shape, and select environments. One needs creative skills and dispositions to generate ideas, analytical skills and dispositions to decide if they are good ideas, and practical skills and dispositions to implement one's ideas and to convince others of their worth.

Because the theory of successful intelligence comprises three subtheories—a componential subtheory dealing with the components of intelligence, an experiential subtheory dealing with the importance of coping with relative novelty and of automatization of information processing, and a contextual subtheory dealing with processes of adaptation, shaping, and selection—the theory has been referred to from time to time as *triarchic.*

Intelligence is not, as Edwin Boring (1923) once suggested, merely what intelligence tests test. Intelligence tests and other tests of cognitive and academic skills measure part of the range of intellectual skills. They do not measure the whole range. One should not conclude that a person who does not test well is not smart. Rather, one should merely look at test scores as one indicator among many of a person's intellectual skills. Moreover, the kinds of skills posited by hierarchical theories (e.g., Carroll 1993; Cattell 1971; Vernon 1971) are viewed only as a subset of the skills important in a broader conception of intelligence.

The Assessment of Successful Intelligence

Our assessments of intelligence have been organized around the analytical, creative, and practical aspects of it. We discuss those assessments here, singly and collectively.

Analytical Intelligence

Analytical intelligence is involved when the information-processing components of intelligence are applied to analyze, evaluate, judge, or compare and contrast. It typically is involved when components are applied to relatively familiar kinds of problems where the judgments to be made are of a fairly abstract nature.

In some early work, it was shown how analytical kinds of problems, such as analogies or syllogisms, can be analyzed componentially (Guyote and Sternberg 1981; Sternberg 1977, 1980b, 1983; Sternberg and Gardner 1983), with response times or error rates decomposed to yield their underlying information-processing components. The goal of this research was to understand the information-processing origins of individual differences in (the analytical aspect of) human intelligence. With componential analysis, one could specify sources of individual differences underlying a factor score such as that for "inductive reasoning." For example, response times on analogies (Sternberg 1977) and linear syllogisms (Sternberg 1980a) were decomposed into their elementary performance components. The general strategy of such research is to (a) specify an information-processing model of task performance; (b) propose a parameterization of this model, so that each information-processing component is assigned a mathematical parameter corresponding to its latency (and another corresponding to its error rate); and (c) construct cognitive tasks administered in such a way that it is possible through mathematical modeling to isolate the parameters of the mathematical model. In this way, it is possible to specify, in the solving of various kinds of problems, several sources of important individual or developmental differences: (1) What performance components are used? (2) How long does it take to execute each component? (3) How susceptible is each component to error? (4) How are the components combined into strategies? (5) What are the mental representations upon which the components act?

Research on the components of human intelligence yielded some interesting results. Consider some examples. First, execution of early components (e.g., inference and mapping) tends exhaustively to consider the attributes of the stimuli, whereas execution of later components (e.g., application) tends to consider the attributes of the stimuli in self-terminating fashion, with only those attributes processed that are essential for reaching a solution (Sternberg 1977). Second, in a study of the development of figural analogical reasoning, it was found that although children generally became quicker in information processing with age, not all components were executed more rapidly with age (Sternberg and Rifkin 1979). The encoding component first showed a decrease in component time with age and then an increase. Apparently, older children realized that their best strategy was to spend more time in encoding the terms of a problem so that they later would be able to spend less time in operating on these encodings. A related, third finding was that better reasoners tend to spend relatively more time than do poorer reasoners in global, up-front metacomponential planning, when they solve difficult reasoning problems. Poorer reasoners, on the other hand, tend to spend relatively more time in local planning (Sternberg 1981). Presumably, the better reasoners recognize that it is better to invest more time up front so as to be able to process a problem more efficiently later on. Fourth, it also was found in verbal analogical reasoning that as children grew older, their strategies shifted so that they relied on word association less and abstract relations more (Sternberg and Nigro 1980).

In the componential analysis work described above, correlations were computed between component scores of individuals and scores on tests of different kinds of psychometric abilities. First, in the studies of inductive reasoning (Sternberg 1977; Sternberg and Gardner 1983), it was found that although inference, mapping, application, comparison, and justification tended to correlate with such tests, the highest correlation typically was with the preparation-response component. This result was puzzling at first, because this component was estimated as the regression constant in the predictive regression equation. This result ended up giving birth to the concept of the metacomponents: higher-order processes used to plan, monitor, and evaluate task performance. Second, it was also found that the correlations obtained for all the components showed convergent-discriminant validation: They tended to be significant with psychometric tests of reasoning but not with psychometric tests of perceptual speed (Sternberg 1977; Sternberg and Gardner 1983). Third, significant correlations with vocabulary tended to be obtained only for encoding of verbal stimuli (Sternberg 1977; Sternberg and Gardner 1983). Fourth, it was found in studies of linear-syllogistic reasoning (e.g., *John is taller than Mary; Mary is taller than Susan; who is tallest?*) that components of the proposed (mixed linguistic-spatial) model that were supposed to correlate with verbal ability did so and did not correlate with spatial ability; components that were supposed to correlate with spatial ability did so and did not correlate with verbal ability. In other words, it was possible successfully to validate the proposed model of linear-syllogistic reasoning not only in terms of the fit of response-time or error data to the predictions of the alternative models but also in terms of the correlations of component scores with psychometric tests of verbal and spatial abilities (Sternberg 1980a). Fifth and finally, it was found that there were individual differences in strategies in solving linear syllogisms, whereby some people used a largely linguistic model, others a largely spatial model, and most the proposed linguistic-spatial mixed model. Thus, sometimes, less than perfect fit of a proposed model to group data may reflect individual differences in strategies among participants.

Creative Intelligence

Intelligence tests contain a range of problems, some of them more novel than others. In some of the componential work, we have shown that when one goes beyond the range of unconventionality of the conventional tests of intelligence, one starts to tap sources of individual differences measured little or not at all by the tests. According to the theory of successful intelligence, (creative) intelligence is particularly well measured by problems assessing how well an individual can cope with relative novelty (see Sternberg et al. 2002).

We presented 80 individuals with novel kinds of reasoning problems that had a single best answer. For example, they might be told that some objects are green and others blue; but still other objects might be grue, meaning green until the year 2000 and blue thereafter, or bleen, meaning blue until the year 2000 and green thereafter. Or they might be told of four kinds of people on the planet Kyron, blens, who are born young and die young; kwefs, who are born old and die old; balts, who are born young and die old; and prosses, who are born old and die young (Sternberg 1982; Tetewsky and Sternberg 1986). Their task was to predict future states from past states, given incomplete information. In another set of studies, 60 people were given more conventional kinds of inductive reasoning problems, such as analogies, series completions, and classifications, but were told to solve them. But the problems had premises preceding them that were either conventional (dancers wear shoes) or novel (dancers eat shoes). The participants had to solve the problems as though the counterfactuals were true (Sternberg and Gastel 1989a, b).

In these studies, we found that correlations with conventional kinds of tests depended on how novel or non-entrenched the conventional tests were. The more novel are the items, the higher are the correlations of our tests with scores on successively more novel conventional tests. Thus, the components isolated for relatively novel items would tend to correlate more highly with more unusual tests of fluid abilities (e.g., that of Cattell and Cattell 1973) than with tests of crystallized abilities. We also found that when response times on the relatively novel problems were componentially analyzed, some components better measured the creative aspect of intelligence than did others. For example, in the "grue-bleen" task mentioned above, the information-processing component requiring people to switch from conventional green-blue thinking to grue-bleen thinking and then back to green-blue thinking again was a particularly good measure of the ability to cope with novelty.

In our original work with divergent reasoning problems having no one best answer, we asked 63 people to create various kinds of products (Lubart and Sternberg 1995; Sternberg and Lubart 1991, 1995, 1996) where an infinite variety of responses were possible. Individuals were asked to create products in the realms of writing, art, advertising, and science. In writing, they were asked to write very short stories for which we would give them a choice of titles, such as "Beyond the Edge" or

"The Octopus's Sneakers." In art, the participants were asked to produce art compositions with titles such as "The Beginning of Time" or "Earth from an Insect's Point of View." In advertising, they were asked to produce advertisements for products such as a brand of bow tie or a brand of doorknob. In science, they were asked to solve problems such as one asking them how people might detect extraterrestrial aliens among us who are seeking to escape detection. Participants created two products in each domain.

First, we found that creativity comprises the components proposed by their investment model of creativity: intelligence, knowledge, thinking styles, personality, and motivation. Second, we found that creativity is relatively although not wholly domain-specific. Correlations of ratings of the creative quality of the products across domains were lower than correlations of ratings and generally were at about the .4 level. Thus, there was some degree of relation across domains at the same time that there was plenty of room for someone to be strong in one or more domains but not in others. Third, we found a range of correlations of measures of creative performance with conventional tests of abilities. As was the case for the correlations obtained with convergent problems, correlations were higher to the extent that problems on the conventional tests were nonentrenched. For example, correlations were higher with fluid than with crystallized ability tests, and correlations were higher, the more novel the fluid test was. These results suggest that tests of creative intelligence have some overlap with conventional tests (e.g., in requiring verbal skills or the ability to analyze one's own ideas—Sternberg and Lubart 1995) but also tap skills beyond those measured even by relatively novel kinds of items on the conventional tests of intelligence.

Practical Intelligence

Practical intelligence involves individuals applying their abilities to the kinds of problems that confront them in daily life, such as on the job or in the home. Practical intelligence involves applying the components of intelligence to experience so as to (a) adapt to, (b) shape, and (c) select environments. People differ in their balance of adaptation, shaping, and selection and in the competence with which they balance among the three possible courses of action.

Much of our work on practical intelligence has centered on the concept of tacit knowledge. We have defined this construct as what one needs to know in order to work effectively in an environment that one is not explicitly taught and that often is not even verbalized (Sternberg et al. 2000; Sternberg and Hedlund 2002; Sternberg and Wagner 1993; Sternberg et al. 1993; Sternberg et al. 1995; Wagner 1987; Wagner and Sternberg 1986; Williams et al. 2002). We represent tacit knowledge in the form of production systems or sequences of "if-then" statements that describe procedures one follows in various kinds of everyday situations.

We typically have measured tacit knowledge using work-related problems that present problems one might encounter on the job. We have measured tacit knowledge for both children and adults, and among adults, for people in over two-dozen occupations, such as management, sales, academia, teaching, school administration, secretarial work, and the military. In a typical tacit-knowledge problem, people are asked to read a story about a problem someone faces and to rate, for each statement in a set of statements, how adequate a solution the statement represents.

In the tacit-knowledge studies, first we have found that practical intelligence as embodied in tacit knowledge increases with experience, but it is profiting from experience, rather than experience per se, that results in increases in scores. Some people could have been in a job for years and still have acquired relatively little tacit knowledge. Second, we also have found that subscores on tests of tacit knowledge—such as for managing oneself, managing others, and managing tasks—correlate significantly with each other. Third, scores on various tests of tacit knowledge, such as for academics and managers, are also correlated fairly substantially (at about the .5 level) with each other. Thus, fourth, tests of tacit knowledge may yield a general factor across these tests.

However, fifth, scores on tacit-knowledge tests do not correlate with scores on conventional tests of intelligence, whether the measures used are single-score measures of multiple-ability batteries. Thus, any general factor from the tacit-knowledge tests is not the same as any general factor from tests of academic abilities (suggesting that neither kind of *g* factor is truly general but rather general only across a limited range of measuring instruments). Sixth, despite the lack of correlation of practical-intellectual with conventional measures, the scores on tacit-knowledge tests predict performance on the job as well as or better than do conventional psychometric intelligence tests. Seventh, in one study done at the Center for Creative Leadership, we further found that scores on our tests of tacit knowledge for management were the best single predictor of performance on a managerial simulation. In a hierarchical regression, scores on conventional tests of intelligence, personality, styles, and interpersonal orientation were entered first and scores on the test of tacit knowledge were entered last. Scores on the test of tacit knowledge were the single best predictor of managerial simulation score. Moreover, these scores also contributed significantly to the prediction even after everything else was entered first into the equation.

Eighth, in work on military leadership (Hedlund et al. 2003; Sternberg and Hedlund 2002; Sternberg et al. 2000), it was found that scores of 562 participants on tests of tacit knowledge for military leadership predicted ratings of leadership effectiveness, whereas scores on a conventional test of intelligence and on a tacit-knowledge test for managers did not significantly predict the ratings of effectiveness. In work with Eskimos (Grigorenko et al. 2004), it was found that low achievers in school can have exceptionally high practical adaptive skills at home.

Even stronger results have been obtained overseas. In a study in Usenge, Kenya, near the town of Kisumu, we were interested in school-age children's ability to adapt to their indigenous environment. We devised a test of practical intelligence for adaptation to the environment (see Sternberg and Grigorenko 1997; Sternberg, Nokes et al. 2001b; and Sternberg 2004, 2007 for more examples of cultural work relevant to the theory). The test of practical intelligence measured children's informal tacit knowledge for natural herbal medicines that the villagers believe can be used to fight various types of infections.

We found no correlation between the test of indigenous tacit knowledge and scores on the fluid-ability tests. But to our surprise, we found statistically significant correlations of the tacit-knowledge tests with the tests of crystallized abilities. The correlations, however, were *negative.* In other words, the higher the children scored on the test of tacit knowledge, the lower they scored, on average, on the tests of crystallized abilities.

We have considered each of the aspects of intelligence separately. How do they fare when they are assessed together?

All Three Aspects of Intelligence Together

Internal Validity Studies Several separate factor-analytic studies support the internal validity of the theory of successful intelligence.

In one study (Sternberg et al. 1999), we used the so-called Sternberg Triarchic Abilities Test (STAT—Sternberg 1993) to investigate the internal validity of the theory. Three hundred twenty-six high-school students, primarily from diverse parts of the United States, took the test, which comprised 12 subtests in all. There were four subtests each measuring analytical, creative, and practical abilities. For each type of ability, there were three multiple-choice tests and one essay test. The multiple-choice tests, in turn, involved, respectively, verbal, quantitative, and figural content.

Confirmatory factor analysis on the data was supportive of the triarchic theory of human intelligence, yielding separate and uncorrelated analytical, creative, and practical factors. The lack of correlation was due to the inclusion of essay as well as multiple-choice subtests. Although multiple-choice tests tended to correlate substantially with multiple-choice tests, their correlations with essay tests were much weaker. The multiple-choice analytical subtest loaded most highly on the analytical factor, but the essay creative and practical subtests loaded most highly on their respective factors. Thus, measurement of creative and practical abilities probably ideally should be accomplished with other kinds of testing instruments that complement multiple-choice instruments.

External Validity Studies We have also looked at the external validity of tests assessing successful intelligence.

The Rainbow Project In a study supported by the College Board (Sternberg and the Rainbow Project Collaborators 2006), we used an expanded set of tests on 1,015 students at 15 different institutions (13 colleges and 2 high schools). Our goal was not to replace the SAT but to devise tests that would supplement the SAT, measuring skills that this test does not measure. In addition to the multiple-choice STAT tests described earlier, we used 3 additional measures of creative skills and 3 of practical skills:

Creative Skills The three additional tests were captioning cartoons, writing creative short stories using two of a number of suggested titles, and orally telling creative stories based on a picture.

Practical Skills The three additional tests were everyday situational judgments based on movie scenarios, a common-sense questionnaire based on problems found in work life, and a common-sense questionnaire based on problems confronted in school.

We found that our tests significantly and substantially improved upon the validity of the SAT for predicting first-year college grades (Sternberg and the Rainbow Project Collaborators 2006). The test also improved equity: using the test to admit a class would result in greater ethnic diversity than would using just the SAT or just the SAT and grade-point average.

The Kaleidoscope Project The Kaleidoscope Project (Sternberg 2005, 2010; Sternberg et al. 2012) has been used over the past 5 years to admit undergraduate students to Tufts University. Each year, all 15,000+ applicants are given a selection of essays assessing analytical, creative, practical, and also wisdom-based skills. The applicants have the option of completing one of the essays, and then the analytical, creative, practical, and wisdom-based skills demonstrated through these essays and other aspects of the application are rated.

The exact Kaleidoscope prompts vary from year to year (see Sternberg 2010 for a complete list through 2009). The questions differ in the skills they emphasize. No question is a "pure" measure of any single component of successful intelligence. Scoring of the exercises is holistic and is completed by admissions officers using rubrics with which they were provided by the Center for the Psychology of Abilities, Competencies, and Expertise at Tufts (PACE Center). We have found that with training, admissions officers can achieve good inter-rater reliability (consistency) in their evaluations.

The results at Tufts illustrated that a highly selective college can introduce an "unconventional" exercise into its undergraduate admissions process without disrupting the quality of the entering class. It is important to underscore the point that academic achievement has always been and remains the most important dimension of Tufts' undergraduate admissions process. Since we introduced the Kaleidoscope pilot in 2006, applications have remained roughly steady or increased slightly, and the mean SAT scores of accepted and enrolling students increased to new highs. In addition, we have not detected statistically meaningful ethnic group differences on the Kaleidoscope measures. Controlling for the academic rating given to applicants by admissions officers (which combines information from the transcript and standardized tests), students rated for Kaleidoscope achieved significantly higher academic averages in their undergraduate work than students who were not so rated by the admissions staff. In addition, research found that students with higher Kaleidoscope ratings were more involved in, and reported getting more out of, extracurricular, active-citizenship and leadership activities in their first year at Tufts.

In sum, as Tufts seeks to identify and develop new leaders for a changing world, Kaleidoscope provides a vehicle to help identify the potential leaders who may be best positioned to make a positive and meaningful difference to the world in the future. In the fast-paced, data-driven atmosphere of highly competitive college admissions, Kaleidoscope validates the role of qualitative measures of student ability and excellence.

Instruction for Successful Intelligence

Instructional studies are a further means of testing the theory (Sternberg & Grigorenko 2007; Sternberg et al. 2001a, 2008, 2009).

Several sets of studies investigated instruction for academic skills. Four sets are briefly described here.

In a first set of studies, researchers explored the question of whether conventional education in school systematically discriminates against children with creative and practical strengths (Sternberg and Clinkenbeard 1995; Sternberg et al. 1996, 1999). Motivating this work was the belief that the systems in most schools strongly tend to favor children with strengths in memory and analytical abilities.

The investigators used the Sternberg Triarchic Abilities Test in some of their instructional work. The test was administered to 326 children around the United States and in some other countries who were identified by their schools as gifted by any standard whatsoever. Children were selected for a summer program in (college-level) psychology if they fell into one of five ability groupings: high analytical, high creative, high practical, high balanced (high in all three abilities), or low balanced (low in all three abilities). Students who came to Yale were then divided into four instructional groups. Students in all four instructional groups used the same introductory-psychology textbook (a preliminary version of Sternberg (1995)) and listened to the same psychology lectures. What differed among them was the type of afternoon discussion section to which they were assigned. They were assigned to an instructional condition that emphasized either memory, analytical, creative, or practical instruction. For example, in the memory condition, they might be asked to describe the main tenets of a major theory of depression. In the analytical condition, they might be asked to compare and contrast two theories of depression. In the creative condition, they might be asked to formulate their own theory of depression. In the practical condition, they might be asked how they could use what they had learned about depression to help a friend who was depressed.

Students in all four instructional conditions were evaluated in terms of their performance on homework, a midterm exam, a final exam, and an independent project. Each type of work was evaluated for memory, analytical, creative, and practical quality. Thus, all students were evaluated in exactly the same way.

Our results suggested the utility of the theory of successful intelligence. This utility showed itself in several ways.

First, we observed when the students arrived at Yale that the students in the high creative and high practical groups were much more diverse in terms of racial, ethnic, socioeconomic, and educational backgrounds than were the students in the high-analytical group, suggesting that correlations of measured intelligence with status variables such as these may be reduced by using a broader conception of intelligence. Thus, the kinds of students identified as strong differed in terms of populations from which they were drawn in comparison with students identified as strong solely by analytical measures. More importantly, just by expanding the range of abilities measured, the investigators discovered intellectual strengths that might not have been apparent through a conventional test.

Second, we found that all three ability tests—analytical, creative, and practical—significantly predicted course performance. When multiple-regression analysis was used, at least two of these ability measures contributed significantly to the prediction of each of the measures of achievement. Perhaps as a reflection of the difficulty of deemphasizing the analytical way of teaching, one of the significant predictors was always the analytical score. (However, in a replication of our study with low-income African-American students from New York, Deborah Coates of the City University of New York found a different pattern of results. Her data indicated that the practical tests were better predictors of course performance than were the analytical measures, suggesting that what ability test predicts what criterion depends on population as well as mode of teaching.)

Third and most importantly, there was an aptitude-treatment interaction whereby students who were placed in instructional conditions that better matched their pattern of abilities outperformed students who were mismatched. In other words, when students are taught in a way that fits how they think, they do better in school. Children with creative and practical abilities, who are almost never taught or assessed in a way that matches their pattern of abilities, may be at a disadvantage in course after course, year after year.

A follow-up study (Sternberg et al. 1998) examined learning of social studies and science by third graders and eighth graders. The 225 third graders were students in a very low-income neighborhood in Raleigh, North Carolina. The 142 eighth graders were students who were largely middle to upper-middle class studying in Baltimore, Maryland, and Fresno, California. In this study, students were assigned to one of three instructional conditions. In the first condition, they were taught the course that basically they would have learned had there been no intervention. The emphasis in the course was on memory. In the second condition, students were taught in a way that emphasized critical (analytical) thinking. In the third condition, they were taught in a way that emphasized analytical, creative, and practical thinking. All students' performance was assessed for memory learning (through multiple-choice assessments) as well as for analytical, creative, and practical learning (through performance assessments).

As expected, students in the successful-intelligence (analytical, creative, practical) condition outperformed the other students in terms of the performance assessments. One could argue that this result merely reflected the way they were taught. Nevertheless, the result suggested that teaching for these kinds of thinking succeeded. More important, however, was the result that children in the successful-intelligence condition outperformed the other children even on the multiple-choice memory tests. In other words, to the extent that one's goal is just to maximize children's memory for information, teaching for successful intelligence is still superior. It enables children to capitalize on their strengths and to correct or to compensate for their weaknesses, and it allows children to encode material in a variety of interesting ways.

We extended these results to reading curricula at the middle-school and the high-school level. In a study of 871 middle-school students and 432 high-school students, we taught reading either triarchically or through the regular curriculum. At the middle-school level, reading was taught explicitly. At the high-school level, reading was infused into instruction in mathematics, physical sciences, social sciences, English, history, foreign languages, and the arts. In all settings, students who were taught triarchically substantially outperformed students who were taught in standard ways (Grigorenko et al. 2002).

The largest-scale study, described in Sternberg et al. (2008), was conducted with 196 teachers and 7,702 students. The study spanned 4 years, 9 states, 14 school districts, and 110 schools. It showed that with many thousands of fourth graders, it was possible to obtain gains in fourth-grade reading and mathematics that were greater for triarchic instruction for critical thinking or memory. This study suggested that triarchic instruction can be "scaled up" to reach children across a wide variety of geographic areas as well as subject matter areas.

Thus, the results of these sets of studies suggest that the theory of successful intelligence is valid as a whole. Moreover, the results suggest that the theory can make a difference not only in laboratory tests but in school classrooms and even the everyday life of adults as well.

Conclusions

This chapter has presented the theory of successful intelligence. Some psychologists believe the theory departs too much from the conventional theory of general intelligence (i.e., the theory of Spearman 1904, 1927). Some disagree with the theory (Gottfredson 2003a, b; Jensen 1998). Others believe the theory does not depart from conventional *g* theory enough (Gardner 1983, 2006). Still others have theories that are more compatible, in spirit, with that proposed here, at least for intelligence (Ceci 1996). The theory is rather newer than that of, say, Spearman (1904) and has much less work to support it as well as a lesser range of empirical support. I doubt the theory is wholly correct—scientific theories so far have not been—but I hope at the same time it serves as a broader basis for future theories than, say, Spearman's theory of general intelligence. No doubt, there will be those who wish to preserve this and related older theories, and those who will continue to do research that replicates hundreds and thousands of time that so-called general intelligence does indeed matter for success in many aspects of life. I agree. At the same time, I suspect it is not sufficient and also that those who keep replicating endlessly the findings of the past are unlikely to serve as the positive intellectual leaders of the future. But only time will tell. As noted earlier, there is typically some value to replication in science, although after a point is established, it seems more to continue to produce papers than to produce new scientific breakthroughs.

The educational system in the United States, as in many other countries, places great emphasis on instruction and assessments that tap into two important skills: memory and analysis. Students who are adept at these two skills tend to profit from the educational system, because the ability tests, instruction, and achievement tests we use all largely measure products and processes emanating from these two kinds of skills. There is a problem, however, namely, that children whose strengths are in other kinds of skills may be shortchanged by this system. These children might learn and test well, if only they were given an opportunity to play to their strengths rather than their weaknesses.

As a society, we can create a closed system that advantages only certain types of children and that disadvantages other types. Children who excel in memory and analytical abilities may end up doing well on ability tests and achievement tests and hence find the doors of opportunity open to them. Children who excel in other abilities may end up doing poorly on the tests and find the doors shut. By treating children with alternative patterns of abilities as losers, we may end up creating harmful self-fulfilling prophecies.

Institutions should consider pooling their resources and developing a common model and common methods of assessment. By working separately, they fail to leverage their strengths and to share information regarding the best ways to make decisions. In essence, each institution "reinvents the wheel." A consortium would be far more powerful than each institution working on its own. Successful intelligence is one model such a consortium might use. Doubtless there are many others. The important thing is to work together toward a common good—toward devising the best ways to select students so as to maximize their positive future impact. We all wish our intellectual leaders to show wisdom. We ourselves need to do the same.

References

Boring, E. G. (1923, June 6). Intelligence as the tests test it. *New Republic*, 35–37.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.

Cattell, R. B., & Cattell, H. E. P. (1973). *Measuring intelligence with the Culture Fair Tests*. Champaign: Institute for Personality and Ability Testing.

Ceci, S. J. (1996). *On intelligence* (Rev. and exp. ed.). Cambridge, MA: Harvard University Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic.

Gardner, H. (2006). *Multiple intelligences: New horizons in theory and practice*. New York: Basic.
Gottfredson, L. S. (2003a). Discussion: On Sternberg's 'Reply to Gottfredson'. *Intelligence, 31*(4), 415–424.
Gottfredson, L. S. (2003b). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence, 31*(4), 343–397.
Grigorenko, E. L., Jarvin, L., & Sternberg, R. J. (2002). School–based tests of the triarchic theory of intelligence: Three settings, three samples, three syllabi. *Contemporary Educational Psychology, 27*, 167–208.
Grigorenko, E. L., Meier, E., Lipka, J., Mohatt, G., Yanez, E., & Sternberg, R. J. (2004). Academic and practical intelligence: A case study of the Yup'ik in Alaska. *Learning and Individual Differences, 14*, 183–207.
Guyote, M. J., & Sternberg, R. J. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology, 13*, 461–525.
Hedlund, J., Forsythe, G. B., Horvath, J. A., Williams, W. M., Snook, S., & Sternberg, R. J. (2003). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *Leadership Quarterly, 14*, 117–140.
(1921). "Intelligence and its measurement": A symposium. *Journal of Educational Psychology, 12*, 123–147, 195–216, 271–275.
Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger/Greenwoood.
Lubart, T. I., & Sternberg, R. J. (1995). An investment approach to creativity: Theory and data. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach*. Cambridge, MA: MIT Press.
Okagaki, L., & Sternberg, R. J. (1993). Parental beliefs and children's school performance. *Child Development, 64*(1), 36–56.
Spearman, C. (1904). 'General intelligence', objectively determined and measured. *American Journal of Psychology, 15*(2), 201–293.
Spearman, C. (1927). *The abilities of man*. London: Macmillan.
Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale: Lawrence Erlbaum Associates.
Sternberg, R. J. (1980a). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109*, 119–159.
Sternberg, R. J. (1980b). Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences, 3*, 573–584.
Sternberg, R. J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology, 73*, 1–16.
Sternberg, R. J. (1982). Natural, unnatural, and supernatural concepts. *Cognitive Psychology, 14*, 451–488.
Sternberg, R. J. (1983). Components of human intelligence. *Cognition, 15*, 1–48.
Sternberg, R. J. (1984). Toward a triarchic theory of human intelligence. *Behavioral and Brain Sciences, 7*, 269–287.
Sternberg, R. J. (1985a). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
Sternberg, R. J. (1985b). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology, 49*(3), 607–627.
Sternberg, R. J. (1993). *Sternberg triarchic abilities test*. Unpublished test.
Sternberg, R. J. (1995). *In search of the human mind*. Orlando: Harcourt Brace College Publishers.
Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.
Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher, 27*, 11–20.
Sternberg, R. J. (1999a). Intelligence as developing expertise. *Contemporary Educational Psychology, 24*, 359–375.
Sternberg, R. J. (1999b). The theory of successful intelligence. *Review of General Psychology, 3*, 292–316.
Sternberg, R. J. (Ed.). (2000). *Handbook of intelligence*. New York: Cambridge University Press.
Sternberg, R. J. (2003). *WICS: A theory of wisdom, intelligence, and creativity, synthesized*. New York: Cambridge University Press.
Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist, 59*(5), 325–338.
Sternberg, R. J. (2005). The theory of successful intelligence. *Interamerican Journal of Psychology, 39*(2), 189–202.
Sternberg, R. J. (2007). Culture, instruction, and assessment. *Comparative Education, 43*(1), 5–22.
Sternberg, R. J. (2008). Schools should nurture wisdom. In B. Z. Presseisen (Ed.), *Teaching for intelligence* (2nd ed., pp. 61–88). Thousand Oaks: Corwin.
Sternberg, R. J. (2009). The Rainbow and Kaleidoscope Projects: A new psychological approach to undergraduate admissions. *European Psychologist, 14*, 279–287.
Sternberg, R. J. (2010). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.
Sternberg, R. J., & Clinkenbeard, P. R. (1995). The triarchic model applied to identifying, teaching, and assessing gifted children. *Roeper Review, 17*(4), 255–260.
Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence?* Norwood: Ablex Publishing Corporation.
Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General, 112*, 80–116.
Sternberg, R. J., & Gastel, J. (1989a). Coping with novelty in human intelligence: An empirical investigation. *Intelligence, 13*, 187–197.
Sternberg, R. J., & Gastel, J. (1989b). If dancers ate their shoes: Inductive reasoning with factual and counterfactual premises. *Memory and Cognition, 17*, 1–10.
Sternberg, R. J., & Grigorenko, E. L. (2007). *Teaching for successful intelligence* (2nd ed.). Thousand Oaks: Corwin.
Sternberg, R. J., & Hedlund, J. (2002). Practical intelligence, g, and work psychology. *Human Performance, 15*(1/2), 143–160.

Sternberg, R. J., & Lubart, T. I. (1991). An investment theory of creativity and its development. *Human Development, 34*(1), 1–31.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist, 51*(7), 677–688.

Sternberg, R. J., & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Development, 51*, 27–38.

Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology, 27*, 195–232.

Sternberg, R. J., & The Rainbow Project Collaborators. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical and creative skills. *Intelligence, 34*(4), 321–350.

Sternberg, R. J., & Wagner, R. K. (1993). The *g*–ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science, 2*(1), 1–4.

Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In H. Reese & J. Puckett (Eds.), *Advances in lifespan development* (pp. 205–227). Hillsdale: Lawrence Erlbaum Associates.

Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist, 50*(11), 912–927.

Sternberg, R. J., Ferrari, M., Clinkenbeard, P. R., & Grigorenko, E. L. (1996). Identification, instruction, and assessment of gifted children: A construct validation of a triarchic model. *Gifted Child Quarterly, 40*, 129–137.

Sternberg, R. J., & Grigorenko, E. L. (1997). The cognitive costs of physical and mental ill health: Applying the psychology of the developed world to the problems of the developing world. *Eye on Psi Chi, 2*(1), 20–27.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998). Teaching triarchically improves school achievement. *Journal of Educational Psychology, 90*, 374–384.

Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude-treatment interaction. *European Journal of Psychological Assessment, 15*(1), 1–11.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J., Snook, S., Williams, W. M., Wagner, R. K., & Grigorenko, E. L. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.

Sternberg, R. J., Grigorenko, E. L., & Jarvin, L. (2001a). Improving reading instruction: The triarchic model. *Educational Leadership, 58*(6), 48–52.

Sternberg, R. J., Nokes, K., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001b). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence, 29*, 401–418.

Sternberg, R. J., Kaufman, J. C., & Pretz, J. E. (2002). *The creativity conundrum: A propulsion model of kinds of creative contributions*. New York: Psychology Press.

Sternberg, R. J., Grigorenko, E. L., & Zhang, L.-F. (2008). Styles of learning and thinking matter in instruction and assessment. *Perspectives on Psychological Science, 3*(6), 486–506.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2009). *Teaching for wisdom, intelligence, creativity, and success*. Thousand Oaks: Corwin.

Sternberg, R. J., Bonney, C. R., Gabora, L., & Merrifield, M. (2012). WICS: A model for college and university admissions. *Educational Psychologist, 47*(1), 30–41.

Tetewsky, S. J., & Sternberg, R. J. (1986). Conceptual and lexical determinants of nonentrenched thinking. *Journal of Memory and Language, 25*, 202–225.

Thurstone, L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Vernon, P. E. (1971). *The structure of human abilities*. London: Methuen.

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond *g:* Putting multiple intelligences theory to the test. *Intelligence, 34*, 487–502.

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology, 52*(6), 1236–1247.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 51–83). New York: Cambridge University Press.

Williams, W. M., Blythe, T., White, N., Li, J., Gardner, H., & Sternberg, R. J. (2002). Practical intelligence for school: Developing metacognitive sources of achievement in adolescence. *Developmental Review, 22*(2), 162–210.

17 Emotional and Social Intelligence and Behavior

Richard E. Boyatzis, James Gaskin, and Hongguo Wei

Of the various levels at which emotional intelligence (EI) and social intelligence (SI) exist within a person, the behavioral level has received the least amount of attention in academic research but holds the most promise for a concept and measurement approach that relates to job and life outcomes, as well as allowing others to “see” EI and SI in action (Goleman 1998; Boyatzis 2009; Cherniss 2010). Emotional intelligence (EI) manifests itself at many levels within a person (Cherniss and Boyatzis 2013). In the past, discussion of EI was often focused on the different theoretical models and different ways of assessing EI (Matthews et al. 2004). In this chapter, we will review the major models or theories which constitute levels of EI and SI and the tests appearing in research publications. The chapter will then focus on the behavioral level and the measurement at this level currently most in use.

R.E. Boyatzis (✉)
Case Western Reserve University, 2077 Kent Road, Cleveland Heights, OH 44106, USA
e-mail: Reb2@case.edu

J. Gaskin
Marriott School of Management, Brigham Young University, Provo, UT, USA
e-mail: James.gaskin@byu.edu

H. Wei
Case Western Reserve University, 1947 E. 120th Street, Apt 3, Cleveland, OH 44106, USA
e-mail: Hongguo.wei@case.edu

A Review of Various Theoretical and Methodological Approaches to EI

To offer a brief review, expanding on Fernandez-Berrocal and Extremera (2006), Boyatzis (2009), and Cherniss (2010), the three major theoretical approaches to EI are described below. Within each theory, we further briefly describe the measures currently appearing in publications. A framework, or grand theory of EI, is offered in Fig. 17.1 to suggest where the ability, self-perception, and behavioral levels coexist (Cherniss and Boyatzis 2013).

Ability-Based Level of EI

Most credit Salovey and Mayer (1990) with the first use of the phrase “emotional intelligence” in a published article. They defined EI as “the subset of social intelligence that involves the ability to monitor one’s own and others’ feelings and emotions, to discriminate among them and to use this information to guide one’s thinking and actions” (Salovey and Mayer 1990). It is based on the view of emotional processes as relying on complex neural activities and as motives to be intelligent (Leeper 1948). It is a type of “emotional information processing” (Salovey and Mayer 1990). They further clarified EI as “the ability to perceive and express emotion, assimilate

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_17,

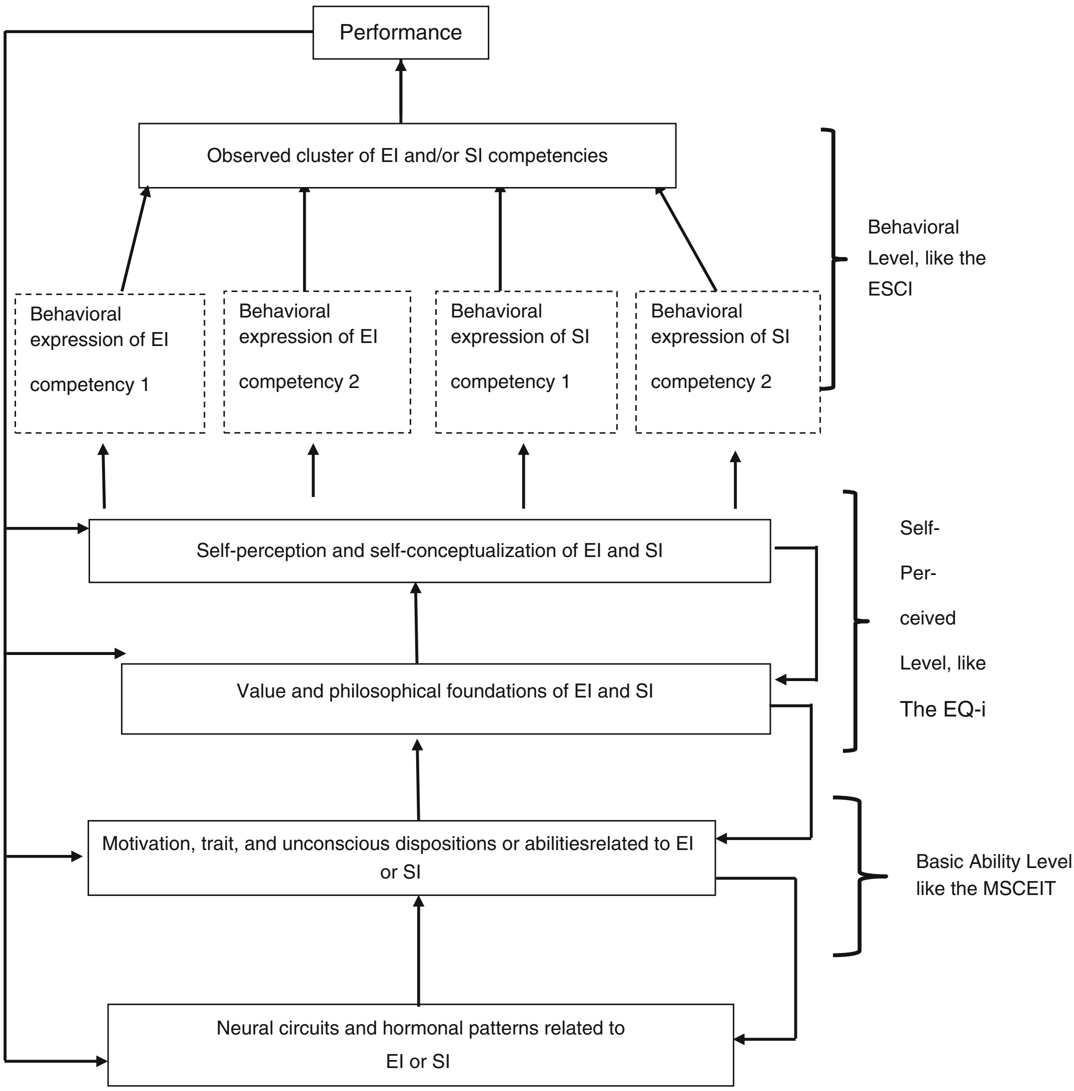


Fig. 17.1 Emotional intelligence (EI), social intelligence (SI), and emotional and social competencies as multiple levels within the personality structure (appearing in Cherniss and Boyatzis (2013) adapted from Boyatzis et al. (2000) and Boyatzis (2009))

emotion in thought, understand and reason." In their theory, Mayer and Salovey (1997) claim that EI is comprised of four dimensions: emotion perception, emotion understanding, emotional facilitation, and emotion regulation.

The measure used most often to examine this approach (EI) is the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). It is a direct performance assessment of emotional processing with some scenarios testing (Mayer et al. 1999; Salovey and Mayer 1997). It is scored in two ways which has caused some confusion: consensus and expert scoring. Because they define EI as a type of cognitive intelligence, it should correlate with cognitive ability (Joseph and Newman 2010). The MSCEIT has content validity (Cherniss 2010). Reliability of the MSCEIT is appropriate, with split-half estimates for the whole scale of .91 and .93 (Mayer et al. 2003). Test-retest reliability has been estimated as $r = .86$

($N=60$) (Brackett and Mayer 2003). The overall internal consistency reliabilities are usually above .75, although the reliabilities of the subscales have not been as good (Conte 2005; Matthews et al. 2002).

A CFA does not support their 4-factor model (Gignac 2005; Rossen et al. 2008). The MSCEIT seems to have discriminant validity with tests of personality. For instance, MSCEIT factors have the strongest correlations with agreeableness from the Big Five ($r=.21$–$.28$), while their correlations with the other four personality factors are less than .20 (Mayer et al. 2008).

The low correlations between MSCEIT factors and relevant constructs impair the convergent validity for the MSCEIT. "Convergent validity for the MSCEIT seems more problematic…. On the other hand, the MSCEIT correlates with measures of verbal intelligence (r=.36) and with other kinds of intelligence (r=.10 to .20) at the levels one would want from a form of intelligence that is supposed to be related to but distinct from other types of intelligence" (Cherniss 2010). For instance, as a form of emotional intelligence measure, MSCEIT is supposed to be moderately correlated but distinct with other emotional perception constructs. However, Roberts et al. (2006) found that the Japanese and Caucasian Brief Affect Recognition Test (JACBART) has low correlations with the overall MSCEIT scale and its subscales ($r=.20$–$.26$) and does not correlate with the emotional perception scale of the MSCEIT at all ($r=.00$). Similarly, the MSCEIT subscales have correlations with the levels of emotional awareness scale (LEAS) from .15 to .20 (Ciarrochi et al. 2003). These results suggest that the MSCEIT has weak or little convergent validity with other relevant emotional constructs.

The incremental validity of the MSCEIT is quantitatively supported. With a student sample, Rossen and Kranzler (2009) found that after controlling for general mental ability and personality, EI measured with the MSCEIT explained incremental variance to positive relationships with others and alcohol use.

Criticism from those other than skeptics of EI in general has focused more on the measure than the concept. That said, the model does seem too restricted to some. Several emotional-related qualities commonly ascribed to EI are excluded, such as "emotional expressiveness, empathy, perspective-taking, self-control, and implicit emotional skills" (Matthews et al. 2006, p. 106). For example, the 4-factor model does not seem plausible. A CFA analysis shows that "the non-constrained 4-factor model yield a non-positive definite matrix, which was interpreted to be due to the fact that two of the branch-level factors (perceiving and facilitating) are collinear" (Gignac 2005, p. 233). Two other studies replicated Gignac's (2005) results (Palmer et al. 2005; Rossen et al. 2008). In addition, the scoring system is complicated (MacCann and Roberts 2008). It is difficult to tell whether a test item is appropriately answered (Matthews et al. 2006). In order to solve this problem, it takes two different approaches in the scoring system: consensus scoring and expert scoring. Although these two scoring approaches are highly correlated ($r=.96$–$.98$), "it is unclear whether a person who thinks about the emotional domain differently from experts or from the average of several peers is low on that ability or whether that person simply has a new (and perhaps better) way of thinking" (Murphy 2006, p. 348).

The format of the MSCEIT has been compared to knowledge tests of EI, which may not provide an appropriate assessment of a person's actual ability (Cherniss 2010). "The assessment of knowledge in the abstract does not reflect the live performance of EI in the rich social situation of real life … one might understand that smiling at someone can be an effective means of producing a positive emotional reaction, but recognizing in a live encounter the moment to smile and doing so in a way that does not seem false or insincere may well be a different ability" (Spector and Johnson 2006, p. 335).

The Schutte Self-Report Emotional Intelligence Test (SREIT) is a self-report scale to measure "a homogeneous construct of emotional intelligence" based on Mayer-Salovey-Caruso model (Schutte et al. 1998). It has 33 items. This test will be reviewed here, but all of the characteristics and criticisms of the self-perception methods described in the next segment apply to this test because the method of assessment is self-report (Schutte et al. 1998).

The SREIT appears to be a face-valid measure of EI (Petrides and Furnham 2000). It has internal consistency (i.e., Cronbach's alpha) of .90 for the 33 items (Schutte et al. 1998). A 2-week test-retest reliability is adequate ($r = .78$) (Conte and Dean 2006). Research results on discriminant validity have been mixed. In an initial small study with 23 college students, the SREIT's correlation with openness to experience was high ($r = .54$), but the absolute correlations with other Big Five personality factors were lower ($r = .21–.28$) (Schutte et al. 1998). However, in one larger study, the SREIT's correlations with the Big Five were from .18 (agreeableness) to .51 (extraversion) (Saklofske et al. 2003). Results from these two studies show that the correlations between SREIT and the Big Five personality factors were not consistent across different samples. In a different study, the correlation between SREIT and psychological well-being was pretty high ($r = .70$) (Brackett and Mayer 2003), which implies that the SREIT is possibly a different measure of psychological well-being.

Since the SREIT is believed to be measuring a type of intelligence, it should have a moderate correlation with the general intelligence (Cherniss 2010). However, a study shows that the SREIT was not significantly correlated with cognitive ability (Saklofske et al. 2003), which suggests that the SREIT did not have the acceptable convergent validity.

The criticisms again as with the MSCEIT focus more on the measure than the concept. For example, it was said that the SREIT cannot measure a general EI factor because the test is not unifactorial (Petrides and Furnham 2000). The 33 items of Schutte's test came from all three subcategories of Salovey and Mayer's original EI model. However, in the analysis, Schutte failed to demonstrate the three subdimensions of EI as separate (Petrides and Furnham 2000). "It would have been more appropriate to perform a factor analysis on the 62 items, rather than a component analysis" (Gignac et al. 2005, p. 1030). Furthermore, it was claimed that the factors should have been rotated obliquely, not orthogonally (Petrides and Furnham 2000).

Another self-assessment test that was based on the MSCEIT model was by Wong and Law which offered "a set of interrelated abilities possessed by individuals to deal with emotions" (Wong and Law 2002; Law et al. 2004). Their construct has four dimensions: self-emotional appraisal, others' emotional appraisal, regulation of emotion, and use of emotion. Their test, the WLEIS, is composed of 16 items. With two samples, the authors showed convergent, discriminant, and incremental validity of this 16-item EI scale (Wong and Law 2002). Their empirical analysis shows that EI has "incremental predictive power over life satisfaction." After controlling for Big Five personality, both the student sample and the work samples showed that others' ratings of EI explained additional variance (Law et al. 2004). All of the criticisms of self-assessment explained in the next section also apply to their test, with the additional major issue that it is only 16 items which dramatically limits the scope of the assessment.

Self-perception Level of EI

Reuven Bar-On developed a model on "an array of non-cognitive capabilities, competencies, and skills that influence one's ability to succeed in coping with environmental demands and pressures" (1997, p. 14). The components included five subtypes of EI: intrapersonal intelligence, interpersonal intelligence, adaptability, stress management, and general mood (Bar-On 1997, 2006). The test, the Emotional Quotient Inventory, is composed of 15 scales in four composites: self-perception includes emotional self-awareness, self-regard, and self-actualization; self-expression includes assertiveness, emotional expression, and independence; interpersonal includes empathy, interpersonal relationships, and social responsibility; decision making includes problem solving, reality testing, and impulse control; stress management includes optimism, flexibility, and stress tolerance. The EQ-i was originally a self-report; in 1997, the 360 version was introduced (Bar-On 1997).

Since the model on which it was based and most, if not all, of the research reported in journals and book chapters to date is from the self-assessment version, it is included in this section. To the extent the published research appears using the 360 version or Bar-On changes his model, it could be included in the subsequent behavioral level section as well.

The EQ-i has good internal consistency. With 243 university students, a study shows that the internal consistency reliability ranged from .69 to .96, with an overall estimate of .96 (Dawda and Hart 2000). Another recent study in North America showed an internal consistency of .97, and a 6-month test-retest reliability of .72 for men ($n=73$) and .80 for women ($n=279$) (Bar-On 2004). However, the internal structure of EQ-i is not consistent. An EFA with a varimax rotation generated a 13-factor model rather than the 15-factor model in Bar-On theory. After removing the problematic factors, a CFA generates a 10-factor model (Bar-On 2006).

Although Bar-On theory was "designed to examine … a conceptual model of emotional and social functioning" (Bar-On 2006, p. 15), the criticism of this approach has been focusing on its overly broad conceptualization (Murphy 2006). Others have claimed that EQ-i has a great deal of overlap with personality, "predictive validity may simply be a consequence of the EQ-i functioning as a proxy measure of personality" (Matthews et al. 2004, p. 16). Thus, the content validity of EQ-i is doubtful given that it includes nonability personality traits and ignores some essential EI factors such as emotional understanding and emotional perception (Cherniss 2010).

The construct validity of Bar-On model and the test confirm that he/she is testing ESC, which is similar to Boyatzis and Goleman model to be explained in the next section as described by Cherniss (2010). An empirical study confirms that EQ-i has convergent validities "with respect to measures of normal personality, depression, somatic symptomatology, intensity of affective experience and alexithymia" (Dawda and Hart 2000, p. 797). Also, the EQ-i has been reported correlating well with some other self-report EI measures ($r=.58$–$.69$) (Bar-On 2004). However, because both convergent and discriminant validity are based on the correlation coefficients, an exceptional good convergent validity is a threat to the discriminant validity. Actually, evidence has shown that the EQ-i does not have good discriminant validity especially considering its relationship with personality variables. Bar-On (2006, p. 16) mentioned that the EQ-i overlaps with personality tests "probably no more than 15 % based on eight studies in which more than 1,700 individuals participated." However, one study shows that when using the Big Five predicts EQ-i scores, the multiple correlation is .79, which means the Big Five personality accounted for the majority of variance in the EQ-i (Grubb and McDaniel 2007).

Another self-assessment test but one based on a composite of the various models of all of the major EI authors is the TEIQue by Petrides and Furnham (2000, 2001, 2003). They claim to assess "trait EI" which they say is "a constellation of emotion-related self-perceived abilities and dispositions located at the lower levels of personality hierarchies" (Petrides and Furnham 2000). It is meant to include all "personality facets that are specifically related to affect" (Petrides et al. 2007, p. 274). The TEIQue has four subdimensions: emotionality, self-control, sociability, and well-being.

After controlling for the Big Five, the TEIQue has a positive relationship with happiness (Furnham and Petrides 2003). It is linked with distinctive reactivity to affect-laden information and has incremental validity over the Big Five (Petrides and Furnham 2003; Petrides et al. 2004). The TEIQue also has incremental validity over alexithymia and optimism (Mikolajczak et al. 2006, 2007).

The criticisms about the TEIQue are similar to those raised about the and self-assessment in general.

The Behavioral Level of EI and SI

The behavioral level has been the most discussed and documented as a "competency" approach (Boyatzis 2009; Cherniss 2010). A competency was originally defined as "an underlying ability that leads to or causes effective performance" (McClelland 1973; Boyatzis 1982; Spencer and Spencer 1993). The various competencies were described as a set of related behavior or actions organized around an underlying intent or context for their use. In this sense, the appearance of the behavior was described as "alternate manifestations" of the underlying characteristics (McClelland 1985).

For example, asking someone questions is an action. But someone can do this for multiple reasons or intent. Someone might ask questions to better understand someone. Another person, or the same person at another time, might ask questions as a way to convince someone else of a position on an issue. The former would result in the combination of intent and action to be labeled or coded as empathy, but the latter would be influence.

One explanation as to why these EI and SI competencies have such a strong relationship to life outcomes and job performance is that they were originally derived inductively by comparing effective and ineffective people in various occupations in a wide variety of organizations in many countries (Boyatzis 1982; Spencer and Spencer 1993; McClelland 1998; Goleman 1998). Outcome criteria or nominations were used to identify samples of people who were effective in a job and those who were not. Detailed collection of their behavior and thoughts using variations of critical incident interviews (Flanagan 1954; Boyatzis 1982; Spencer and Spencer 1993) or videotaped simulations enabled researchers to develop "codes" to determine whether or not a person was demonstrating these EI and SI competencies in various situations. For more detailed explanation of the methods, see Boyatzis (1982, 1998) or Spencer and Spencer (1993), or Boyatzis (2009) for a review of the methods.

Once these "codes" of competencies were applied in many settings and jobs, 360 assessments were developed to collect the observations from a broader array of sources and make the data collection easier (Boyatzis 2009). A 360 measurement is a multisource assessment in which people who live and work with a target person describe that person's typical behavior through a series of questions. Items for the 360 were created based on the behavior validated in the earlier studies with a phrase added that provided the respondent with the intent. For example, the behavior of "listens attentively" was validated as an indicator of empathy in many of the competency studies. To include it in the 360, we added the intent. The item now reads, "Understands others by listening attentively."

The most widely used and cited of the behavioral level approaches to EI comes from McClelland, Boyatzis, Goleman, Cherniss, and their colleagues at The Hay Group. Several definitions have appeared over the decades. One was that "Emotional intelligence [includes] abilities such as being able to motivate oneself and persist in the face of frustrations; to control impulse and delay gratification; to regulate one's moods and keep distress from swamping the ability to think; to empathize and to hope" (Goleman 1995, p. 34). In describing their definitions of these clusters in detail, Boyatzis (2009) said, "(a) an emotional, intelligence competency is an ability to recognize, understand, and use emotional information about oneself that leads to or causes effective or superior performance; (b) a social intelligence competency is the ability to recognize, understand and use emotional information about others that leads to or causes effective or superior performance; and (c) a cognitive intelligence competency is an ability to think or analyze information and situations that leads to or causes effective or superior performance."

In the Boyatzis and Goleman model, EI has two dimensions and SI has two dimensions: EI includes self-awareness and self-management and SI includes social awareness and relationship management. The Emotional and Social Competency Inventory (ESCI), or its university version (ESCI-U which includes two cognitive competencies), is more outcome oriented because of the way in which they were developed (Boyatzis and Goleman 1996, 1999; Wolff 2007, 2008; Hay Group 2011). Detailed statistical analysis and results about the psychometrics of the tests are summarized in the following section.

The ESCI assesses 12 competencies and the ESCI-U assesses 14 competencies. Past research has shown a set of competencies that differentiate outstanding and effective performers from others in many countries (Bray et al. 1974; Boyatzis 1982; Kotter 1982; Thornton and Byham 1982; Luthans et al. 1988; Howard and Bray 1988; Campbell et al. 1970; Spencer and Spencer 1993; Goleman 1998; Hopkins and Bilimoria 2008; Koman and Wolff 2008; Dreyfus 2008; Williams 2008; Sternberg 1996). They can be clustered into three sets of competencies, two of which are related to EI and SI: (1) cognitive competencies, such as systems thinking and pattern recognition; (2) emotional intelligence competencies, including self-awareness and self-management competencies, such as emotional self-awareness and emotional self-control, and what used to be a part of EI but is now separated for theoretical as well as empirical reasons; and (3) social intelligence competencies, including social awareness and relationship management competencies, such as empathy and teamwork. The specific competencies considered to be a behavioral approach to emotional, social, and cognitive intelligence are shown in Table 17.1.

The criticisms of this approach are that the concepts are over-inclusive (Matthews et al. 2004). Other criticisms of Matthews et al. (2004) have been addressed in publications in the last 8 years and are reviewed in this section of this chapter.

There has been one other measure appearing in selected publications which was developed to assess the behavioral level of EI. Dulewicz and Higgs (1999) developed a measure based on the work of Boyatzis (1982) and Boyatzis and Goleman (1996). They assessed being aware of and managing one's own feelings and emotions, being sensitive to and influencing others, sustaining one's motivation, and balancing one's motivation and drive with intuitive, conscientious, and ethical behavior (Dulewicz and Higgs 1999). Their measure, the EIQ, has seven dimensions: self-awareness, emotional resilience, motivation, interpersonal sensitivity, influence, intuitiveness, and conscientiousness (Higgs and Dulewicz 2002; Dulewicz et al. 2003). It is also a 360. The alpha for overall EIQ was .77. The alpha coefficients for each of the element scales ranged from .6 to .8 (Dulewicz and Higgs 2000). The EIQ was significantly related with performance measures (except for interpersonal sensitivity)

Table 17.1 The scales and clusters of the Emotional and Social Competency Inventory (ESCI and ESCI-U—the university version) (Boyatzis and Goleman 1996, 1999; Boyatzis et al. 2001, 2007)

Emotional intelligence competencies are as follows:
***Self-awareness cluster*: Concerns knowing one's internal states, preferences, resources, and intuitions. The self-awareness cluster contains one competency:
Emotional self-awareness: Recognizing one's emotions and their effects
Self-management cluster: Refers to managing one's internal states, impulses, and resources. The self-management cluster contains four competencies:
Emotional self-control: Keeping disruptive emotions and impulses in check
Adaptability: Flexibility in handling change
Achievement orientation: Striving to improve or meeting a standard of excellence
Positive outlook: Seeing the positive aspects of things and the future
Social intelligence competencies are as follows:
Social awareness cluster: Refers to how people handle relationships and awareness of others' feelings, needs, and concerns. The social awareness cluster contains two competencies:
Empathy: Sensing others' feelings and perspectives and taking an active interest in their concerns
Organizational awareness: Reading a group's emotional currents and power relationships
Relationship management cluster: Concerns the skill or adeptness at inducing desirable responses in others. The cluster contains five competencies:
Coach and mentor: Sensing others' development needs and bolstering their abilities
Inspirational leadership: Inspiring and guiding individuals and groups
Influence: Wielding effective tactics for persuasion
Conflict management: Negotiating and resolving disagreements
Teamwork: Working with others toward shared goals. Creating group synergy in pursuing collective goals
Cognitive intelligence competencies (in the ESCI university version only) are as follows:
Systems thinking: Perceiving multiple causal relationships in understanding phenomena or events
Pattern recognition: Perceiving themes or patterns in seemingly random items, events, or phenomena

(Dulewicz and Higgs 2000). Dulewicz et al. (2003) tested EIQ's content, construct, and criterion-related validity and showed support.

The ESCI and ESCI-U

The ESCI is a 68-item test assessing the 12 EI and SI competencies listed in Table 17.1. The ESCI-U (university version) is a 70-item test assessing the 12 EI and SI competencies in Table 17.1, as well as the two CI (cognitive) competencies listed. The tests were designed to be used as 360, multisource rater instruments. In most research, only the "other" assessments from informants are used. The self-assessment is used primarily in applications in coaching, training, and college courses, along with the "other" assessments. The following statistical tests were computed on the ESCI with a sample of 5,761 self-assessments and 62,297 other assessments. The effective sample sizes for CFAs with no missing data were 4,468 and 25,057, respectively. The sample was generated from research studies and training programs with managers and professionals from many countries of the world. The following statistical tests were computed on the ESCI-U with a sample of 1,629 self-assessments and 21,288 other assessments. The effective sample sizes for CFAs with no missing data were 1,398 and 8,981, respectively. The ESCI-U sample was generated from research and graduate courses with MBAs from the Weatherhead School of Management at Case Western Reserve University.

Scale Reliabilities Using Alpha

Cronbach's alpha was computed for each scale shown in Table 17.2. As a comparison, the earlier version of the test, called the ECI-2 "others," showed an average alpha of 0.78 and since 2007 the ESCI "others" shows 0.87. Thus, the ESCI shows improved scale reliability as well as better factor structure. To support aggregation across various informant sources, inter-rater agreement scores (rWG(J)) were calculated. Badri (2013), using a sample of 468 respondents, reported the following rWG(J) for each competency: emotional self-awareness, 0.86; achievement orientation, 0.94; adaptability, 0.95; emotional self-control, 0.92; positive outlook, 0.95; empathy, 0.92; organizational awareness, 0.92; coach and mentor, 0.91; influence, 0.93; conflict management, 0.88; inspirational leadership, 0.94; and teamwork, 0.97.

Model Fit Using CFA

An exploratory factor analyses on each of the four data sets (ESCI self, ESCI others, ESCI-U self, and ESCI-U others) showed almost all of the items loaded on the appropriate, predicted scale (Boyatzis and Gaskin 2010; Hay Group 2011). Nonconforming items were dropped from the analysis. Several items which were close to the typical cutoffs were slightly reworded for future use of the test but kept in the subsequent analyses reported here. Confirmatory factor analyses were run on each of the four samples to assess the model fit to the 12 scales in the ESCI and 14 scales in the ESCI-U.

Except for the chi-square to df ratio for the ESCI self and other and ESCI-U other, all estimates of model fit were within acceptable standards to support the scales as an acceptable model. In the chi-square to df ratio case, it is believed that the large sample size contributed to an inflation of the chi-square, thus rendering that estimate less useful. Specifically, for the self and the other, independently calculated, for the ESCI and ESCI-U, the RMSEA, PCLOSE, CFI, PCFI, GFI, and SRMR are within desired levels for a satisfactory model fit. Prior to running the CFA, all subjects with missing data were eliminated, resulting in the smaller sample sizes reported in Table 17.3.

Convergent and Discriminant Validity of the Scales

PLS-GRAPH version 3.0 was used to determine the reliability and validity of the scales. Partial

Table 17.2 Cronbach's alpha for each scale in the ESCI and ESCI-U (sample sizes are shown in parentheses following the alpha)

	ESCI		ESCI-U	
	Self	Other	Self	Other
Emotional self-awareness	.754 (5,534)	.827 (42,215)	.771 (1,605)	.804 (16,223)
Emotional self-control	.854 (5,664)	.910 (56,713)	.843 (1,611)	.882 (17,993)
Achievement orientation	.800 (5,668)	.861 (47,975)	.705 (1,621)	.779 (18,381)
Adaptability	.720 (5,573)	.845 (53,875)	.752 (1,605)	.820 (16,914)
Positive outlook	.829 (5,641)	.883 (54,598)	.825 (1,606)	.846 (17,164)
Empathy	.708 (5,638)	.856 (52,138)	.725 (1,622)	.836 (17,520)
Organizational awareness	.786 (5,579)	.861 (54,475)	.764 (1,603)	.830 (17,159)
Influence	.721 (5,606)	.835 (50,702)	.746 (1,569)	.822 (14,693)
Teamwork	.771 (5,668)	.886 (58,911)	.775 (1,616)	.857 (18,354)
Coach and mentor	.856 (5,546)	.920 (48,744)	.821 (1,590)	.868 (15,369)
Conflict management	.682 (5,607)	.785 (51,948)	.636 (1,592)	.733 (16,252)
Inspirational leadership	.897 (5,221)	.887 (51,199)	.817 (1,536)	.866 (15,749)
Systems thinking	NA	NA	.794 (1,582)	.821 (15,176)
Pattern recognition	NA	NA	.792 (1,577)	.831 (13,502)

Table 17.3 Model fit using CFA using AMOS for the ESCI and ESCI-U both self and other responses for each

	ESCI		ESCI-U	
	Self	Other	Self	Other
Effective *n*	4,468	25,057	1,398	8,981
Chi-square	18,793	100,823	6,606	31,746
df	2,013	2,013	1,921	1,921
RMSEA	.043	.044	.042	.042
PCLOSE	1.000	1.000	1.000	1.000
CFI	.849	.910	.875	.919
PCFI	.797	.855	.808	.849
GFI	.861	.860	.862	.888
SRMR	.0536	.0424	.0469	.0319

least squares was used rather than the traditional covariance-based SEM tools such as LISREL or AMOS. PLS is especially well suited for working with new measures (Chin et al. 2003).

To test for convergent validity, items in each construct must have loadings over 0.5 (Fornell 1982; Hair et al. 1995) and composite reliabilities (CR) should be over 0.7 (Nunnally et al. 1994) and greater than their respective average variance extracted (AVE). Lastly, the average variance extracted should be maximized, with a minimum of 50 % (Barclay et al. 1995). Discriminant validity is established by showing that the correlation between any two constructs is less than the square root of the average variance extracted by the measures of that construct (Lim et al. 2006).[1]

As shown in Tables 17.4 through 17.7, for each of the scales in each of the analyses, all construct composite reliabilities, average variance extracted, and their relationship have met the respective thresholds to be considered sufficiently convergent. In addition, the square root of the average variance extracted was greater than the inter-scale correlations, showing discriminant validity for each of the scales for each of the analyses. These indicators of convergent and discriminant validity are shown for the ESCI self in Table 17.4, the ESCI other in Table 17.5, the ESCI-U self in Table 17.6, and for the ESCI-U

[1]Almost no one reports maximum shared variance any longer as an indication of discriminant validity, but for those aficionados who still use it, we have a different position. In many analyses of this type, another measure is also recommended for showing discriminant validity, called the maximum shared variance (MSV). But in this case, in which the theoretical model on which the items and scales were built is a circumplex model, it is assumed that some items, as well as scales, will have a high shared variance with others. For example, an expression or use of empathy is required to show inspirational leadership. As a result, the maximum shared variance of an item or scale will be deceptive and not an appropriate indicator of discriminant or discriminant validity.

Table 17.4 ESCI self: Convergent and discriminant validity using PLS ($n=4{,}468$) [*square root of the AVE on the diagonal*]

Scale	μ	SD	CR	AVE	ESA	AO	ESC	A	PO	EMP	OA	IL	CFM	CM	TEAM
ESA	3.744	.817	.848	.521	**.722**										
AO	4.278	.695	.861	.554	.314	**.744**									
ESC	4.023	.698	.895	.631	.175	.221	**.794**								
A	4.127	.687	.826	.543	.301	.456	.368	**.737**							
PO	4.182	.683	.875	.540	.279	.409	.419	.427	**.735**						
EMP	3.983	.663	.819	.532	.448	.296	.441	.406	.380	**.729**					
OA	4.243	.670	.860	.551	.415	.319	.321	.474	.367	.482	**.742**				
IL	3.955	.743	.863	.652	.391	.483	.288	.497	.533	.428	.452	**.807**			
CFM	3.867	.745	.882	.627	.365	.369	.196	.344	.327	.357	.342	.445	**.792**		
CM	4.130	.747	.902	.697	.341	.447	.222	.350	.383	.379	.362	.598	.372	**.835**	
TEAM	4.338	.642	.856	.544	.355	.347	.407	.347	.420	.545	.455	.519	.323	.458	**.737**
INF	3.848	.773	.801	.575	.290	.280	.212	.378	.325	.320	.435	.413	.252	.363	.358

Table 17.5 ESCI Other: Convergent and discriminant validity using PLS convergent and discriminant validity using PLS ($n = 12{,}419$) [*square root of the AVE on the diagonal*]

Scale	μ	SD	CR	AVE	ESA	AO	ESC	A	PO	EMP	OA	IL	CFM	CM	**TEAM**
ESA	4.136	.795	.879	.549	**.741**										
AO	4.315	.768	.903	.608	.570	**.780**									
ESC	3.892	.844	.932	.697	.496	.498	**.835**								
A	4.058	.877	.891	.578	.586	.697	.590	**.760**							
PO	3.880	.852	.919	.654	.563	.626	.604	.685	**.808**						
EMP	3.806	.841	.889	.667	.711	.573	.637	.655	.623	**.817**					
OA	4.124	.815	.901	.645	.608	.579	.546	.705	.597	.700	**.803**				
IL	4.036	.850	.918	.692	.639	.714	.547	.734	.686	.710	.689	**.832**			
CFM	4.024	.889	.865	.617	.617	.601	.546	.648	.593	.668	.601	.689	**.786**		
CM	4.298	.758	.939	.720	.611	.678	.499	.635	.588	.674	.622	.796	.638	**.848**	
TEAM	4.192	.755	.920	.658	.650	.645	.659	.670	.674	.775	.692	.766	.667	.727	**.811**
INF	4.262	.802	.871	.575	.604	.582	.520	.678	.626	.656	.693	.688	.598	.629	.687

Table 17.6 ESCI-U self: Convergent and discriminant validity using PLS [*square root of the AVE on the diagonal*] ($n = 1{,}398$)

Scale	μ	SD	CR	AVE	A	AO	CFM	CM	EMP	ESA	ESC	INF	IL	OA	PO	PR	ST
A	4.010	.751	.850	.531	**.729**												
AO	4.159	.770	.821	.534	.592	**.731**											
CFM	3.678	.874	.880	.591	.493	.449	**.769**										
CM	3.717	.890	.884	.606	.407	.429	.482	**.779**									
EMP	3.945	.771	.816	.527	.449	.353	.446	.384	**.726**								
ESA	3.889	.839	.845	.525	.443	.453	.554	.445	.482	**.724**							
ESC	3.892	.827	.893	.677	.565	.345	.423	.309	.454	.326	**.823**						
INF	3.797	.810	.842	.516	.584	.546	.563	.525	.458	.526	.378	**.718**					
IL	3.675	.864	.878	.591	.527	.534	.540	.657	.428	.484	.369	.649	**.769**				
OA	4.118	.735	.855	.541	.547	.439	.455	.429	.489	.491	.383	.631	.546	**.736**			
PO	4.023	.821	.878	.591	.517	.510	.495	.463	.376	.445	.414	.448	.564	.410	**.768**		
PR	3.785	.876	.863	.559	.485	.431	.421	.339	.375	.455	.343	.536	.431	.393	.333	**.747**	
ST	3.841	.787	.871	.629	.544	.490	.418	.391	.456	.468	.367	.534	.455	.446	.359	.674	**.793**
TEAM	4.118	.742	.850	.533	.471	.440	.468	.525	.514	.460	.433	.510	.589	.550	.457	.339	.424

Table 17.7 ESCI-U other: Convergent and discriminant validity using PLS [*square root of the AVE on the diagonal*] ($n = 8,981$)

Scale	μ	SD	CR	AVE	A	AO	CFM	CM	EMP	ESA	ESC	INF	IL	OA	PO	PR	ST
A	4.192	.776	.884	.604	**.777**												
AO	4.330	.753	.871	.628	.747	**.792**											
CFM	3.969	.861	.911	.663	.640	.580	**.814**										
CM	4.079	.875	.906	.659	.655	.658	.658	**.812**									
EMP	4.034	.838	.885	.657	.677	.602	.671	.671	**.811**								
ESA	4.031	.841	.875	.585	.643	.607	.694	.634	.708	**.765**							
ESC	4.134	.820	.906	.707	.682	.575	.597	.555	.674	.598	**.841**						
INF	4.033	.816	.884	.603	.696	.657	.725	.700	.697	.693	.572	**.776**					
IL	4.032	.871	.906	.658	.695	.673	.691	.793	.704	.668	.595	.758	**.811**				
OA	4.290	.751	.886	.608	.742	.691	.623	.659	.692	.652	.625	.724	.690	**.780**			
PO	4.241	.769	.899	.639	.704	.694	.645	.657	.658	.660	.630	.675	.711	.669	**.800**		
PR	3.973	.835	.890	.618	.679	.623	.647	.621	.652	.656	.553	.729	.676	.648	.613	**.786**	
ST	4.046	.792	.885	.658	.714	.647	.641	.647	.667	.663	.591	.696	.672	.672	.616	.773	**.811**
TEAM	4.278	.776	.907	.661	.701	.687	.630	.730	.736	.642	.650	.678	.758	.731	.704	.575	.614

other in Table 17.7. For the ESCI other analysis, the sample size exceeded available memory, so half of the sample was randomly chosen for this analysis. In the ESCI self, some items were dropped from to conform to desirable loadings. The specific items dropped for the ESCI self analysis included seven reverse-scored items (reverse-scored items for emotional self-control, coaching and mentoring, achievement orientation, inspirational leadership, teamwork, adaptability, and empathy). For this same analysis, six other items were dropped: one from emotional self-awareness, one from coaching and mentoring, one from adaptability, one from conflict management, and one from influence. Each of these items has been reworded to possibly correct for the loadings in the latest revision to the ESCI. For the ESCI other analysis, only one empathy item was dropped. For research purposes, a list of these specific items is available from the first author. No items had to be trimmed for the ESCI-U self or other.

Validity of the Behavioral Level with Coded Interviews

One form of assessment of the behavioral level of EI and SI is the critical incident interview, also known as the behavioral event interview (Flanagan 1954; Boyatzis 1982; Spencer and Spencer 1993). Work samples are collected and then coded by blind coders with high inter-rater reliability (Boyatzis 1998). Boyatzis (1982) provided the first published validation of these competencies against effectiveness measures in a sample of 253 managers and executives from six large private sector organizations and six large government organizations which included samples from the US Navy and Marine Corps. Although many studies followed in the 1980s and 1990s, these studies were often done by consultants and were not published.

A series of studies done in the 1990s were published later, showing the link between the use of the EI and SI competencies and work effectiveness using the critical incident interviews. Dreyfus (2008) showed that the EI and SI competencies predicted effectiveness of branch chiefs in research (middle-level managers) at NASA. Williams (2008) showed they predicted effectiveness of elementary, middle, and high school principals in a large urban city. Boyatzis and Ratti (2009) made the same links in samples from a large company in Italy and top executives from cooperatives also in Italy. Ryan et al. (2009) reported a significant prediction of effectiveness of the EI and SI competencies in executives in a variety of European companies. Gutierrez et al. (2012) showed the links to effectiveness of coded EI and SI competencies in top executives in India and China and compared them to comparable samples from Western Europe.

Another benefit of assessing EI and SI at the behavioral level is that it should be more amenable to change and improvement than other levels. To document change, EI and SI competencies were shown to significantly improve for four cohorts of full-time MBA and two cohorts of part-time MBA students as compared to two cohorts of each type of MBA in a baseline study in earlier years (Boyatzis et al. 2002) in the USA and in Italy (Camuffo et al. 2009).

Validity of the Behavioral Level with Early Versions of the ECI or ECI-U

Using a 360 or multisource assessment as a way to measure EI and SI in behavior—as seen by others—provides an easier assessment tool. It is less costly in human effort to collect and score and provides more of a consensual validation that improves face validity of the results. It is a test of the behavioral level of EI and SI.

Using a company-customized variation of the earlier version of the ESCI, called the SAQ/EAQ (Boyatzis 2009), Boyatzis (2006) conducted a tipping point analysis and showed the operating profit contribution of senior partners per year in a longitudinal study of executive performance in a major international consulting company. When senior partners used EI competencies above the tipping point (as seen by others), as compared to those senior partners showing the competencies below the tipping point, they showed 78 % high gross operating profit from their accounts for the EI cluster of self-management and 390 % increase for the EI cluster which they called self-regulation. In the same study, the increase in operating profit from senior partners above the tipping point (again, as seen by others) in SI competencies was 110 %. Even the operating profit contribution of senior partners was 50 % for those above the tipping point in demonstrated cognitive intelligence competencies.

Using the earlier version of the ESCI, called the ECI-2, Hopkins and Bilimoria (2008) showed that female bank executives showed that the use of the EI and SI competencies (as seen by others) was associated with higher measures of success in terms of the company's performance appraisal system. Koman and Wolff (2008) showed a similar pattern for US Navy Commanding Officers of flight crews when assessed against actual performance of the flight in combat simulations and air maneuvers. EI/SI as seen by others was significantly related to job performance at Johnson & Johnson in a study of 358 managers (Cavallo and Brienza 2002).

In another domain, Boyatzis et al. (2011) showed that Catholic Pastor priests who demonstrated more EI and SI competencies as seen by others working with them showed higher parishioner satisfaction on a complex measure of eight dimensions of parishioner involvement in the church and community. It did not predict church attendance nor donations per family.

In a study of Spanish nonprofit executives, Ramo et al. (2009) showed that using more EI and SI competencies predicted effectiveness. Camuffo et al. (2012) showed a similar set of findings for Italian executives from a variety of companies in Northern Italy. (Nel 2001; Aliaga and Taylor 2012) showed that using more EI and SI resulted in more effectiveness in Peruvian managers in a copper refinery. Sharma (20121) showed that Indian middle-level managers from large companies and public sector organizations who demonstrated more EI and SI than others were more effective. Sevinc (2001) reported EI/SI scores (as seen by others) were significantly correlated with salary, job, and life satisfaction of 71 Turkish professionals in the financial services sector. In South African call centers, the EI/Si as seen by others of 135 call agents was significantly related to job performance (Nel 2001). A study of 33 business development managers at Bass Brewers in the UK showed that EI/SI, as seen by others, was significantly predictive of a multidimensional measure of job performance (Lloyd 2001). All of these studies used the ECI-2.

Van Oosten (2013) showed that demonstrating more EI and SI competencies, as measured with the ECI-2, predicted performance of bank executives. Her findings suggested that certain of these competencies, the ones associated with leading change were the most effective. In a study of first responders, fire fighters and officers in London, Stagg and Gunter (2002) showed that EI/SI (again, as seen by others) was statistically related to a battery of performance measures.

In comparing performance against a traditional measure of academic cognitive intelligence, Victoroff and Boyatzis (2013) showed that assessment of the EI and SI competencies, as seen by other dental graduate students 6 weeks into their program, predicted their grades in the third and fourth year courses, which all take place in the dental clinic working with patients under faculty supervision. Meanwhile, the Dental Admissions Test (DAT) predicted grades in the first and second years of the graduate dental program which are all didactic courses, but showed no prediction of their grades in their third and fourth clinical years.

This result supports the finding of discriminant validity of EI/SI competencies with traditional measures of cognitive intelligence. Murensky (2000) studied 90 oil company executives and found no relationship between Watson-Glaser Test of Critical Thinking and the EI and SI competencies. She found three of the SI competencies to have a negative correlation to the Watson-Glaser. In Philippine plants of two multinationals, Sergio (2001) showed that others' assessments of 134 plant supervisors' EI/SI was significantly linked to job performance. A mental ability test was also linked to performance, but the mental ability test and EI/SI were not correlated. These studies support the idea that behavioral EI/SI is a different human capability or talent than cognitive intelligence.

Again, assessment at this level of EI and SI allows for careful studies of change. Boyatzis et al. (2002) reported on one cohort of graduating full-time MBAs with the SAQ/EAQ; Boyatzis and Saatcioglu (2008) were able to add another cohort of full-time MBAs using the SAQ/EAQ and two more cohorts using the ECI-U (the earlier version of the ESCI-U); and Boyatzis et al. (2010) added yet another cohort of full-time MBAs with the ECI-U showing dramatic improvement on the EI and SI competencies, as compared to baseline years and the MBAs themselves at entry into the program.

Validity of the Behavioral Level with the ESCI or ESCI-U

The ESCI and ESCI-U reflected several major improvements on the earlier versions of the test of EI and SI at the behavioral level. First launched in 2007, the validation research is just emerging. Studying the effectiveness of sales executives in a financial services company, Boyatzis et al. (2012) found that demonstrated EI and SI predicted effectiveness. They were able to assess cognitive intelligence using Ravens Progressive Matrices and personality using the NEO-PR and found that EI/SI competencies as seen by others predicted effectiveness, but cognitive g and personality did not. In addition, they were able to focus, through multiple regressions, on the fact that adaptability was the most powerful predictor of leader effectiveness among the EI competencies and influence was among the SI competencies.

In a study of leaders from a variety of companies, Havers (2010) showed "leaders who demonstrated fewer than three ESCI strengths drew upon a limited range of leadership styles, tending to rely primarily on the coercive style … In contrast, leaders with 10 or more ESCI strengths used a much wider range of leadership styles, including those likely to engage their team members, providing long term direction and vision, creating harmony, encouraging new ideas and investing in others development" (p. 2). Furthermore, Havers (2010) reported that 92 % of the leaders showing high emotional self-awareness created a positive organizational climate (as seen by their subordinates), while 78 % of the leaders with low emotional self-awareness created a negative climate.

Also in the sales function, Lisicki (2011) showed that EI and SI competencies predicted sales performance of sales people in a pharmaceutical company. He showed that coaching and mentoring competency was the most potent factor in predicting sales performance.

A leader's EI and SI, as assessed with the ESCI, is associated with their subordinates' job performance and satisfaction in a study of 20 leaders form a multinational in Egypt (Shams 2008). She found that subordinates' job satisfaction was significantly correlated with all of the EI competencies and four of the seven SI competencies. Meanwhile, subordinates' job performance was correlated with emotional self-awareness and emotional self-control from the EI cluster and influence from the SI cluster.

In a sample of Indian managers, Badri (2013) showed that EI and SI competencies as seen by others assessed with the ESCI predicted leader effectiveness and use of the transformational leadership style. He found a relationship to transactional leadership style as well, but only for one subscale, and others were negatively related to use of EI and SI. This supported the findings of Piel (2008) in a study of 82 project managers from various companies that EI and SI, as assessed with the ESCI, was related to using the transformational leadership style.

Using the self-assessment from the ESCI, Quinn (2013) showed that it predicted physician leader effectiveness in terms of participation as a leader in hospitals. Pittenger (2012) showed that for IT managers and advanced professionals from a variety of companies, EI and SI competencies predicted effectiveness in terms of engagement.

Using a 360 from their own competency model of EI and SI, Young and Dulewicz (2009) showed the same pattern found earlier with the US Navy in that EI and SI as seen by others predicted leader effectiveness for British Naval officers. Ryan et al. (2012) showed that EI and SI with a 360 from a customized competency assessment in a large Swiss company predicted executive effectiveness.

Conclusion

The behavioral level does provide theoretical and empirical support for the relevance of EI and SI in predicting effectiveness and work and life outcomes. It also shows that at least one measure of the behavioral level, the ESCI or ESCI-U, satisfies appropriate psychometric standards of a reliable and valid test, with appropriate convergent and discriminant validity, and is a good model fit for each of the separate scales.

This explanation of the levels of EI may serve to refocus the intellectual debate on the conditions when the relationship of EI (and SI) helps us to understand performance, effectiveness, and life and job outcomes. It offers a theoretical rationale as to how the three approaches to EI do function in unison within the person. But the different levels require different types of measurement. This is where the ESCI can complement the MSCEIT and EQ-i in research.

Practical implications of the behavioral level are considerable. It is far easier to document improvement on the behavioral level than other levels as has been shown in published research. The behavioral level guides coaches and trainers seeking to improve performance of executives, managers, and leaders, along with professionals in organizations. It can be a guide and tool for graduate and undergraduate programs seeking to develop the whole person, not just their knowledge, and help such academic programs with the needed outcome assessment for accreditation and documentation of program impact.

References

Aliaga Araujo, S. V., & Taylor, S. N. (2012). The influence of emotional and social competencies on the performance of Peruvian refinery staff. *Cross Cultural Management, 19*(1), 19–29.

Badri, N. (2013). *An examination of the relationship between emotional intelligence leadership styles and leadership effectiveness*. Unpublished doctoral dissertation, Jaypee Institute of Information Technology.

Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to causal modeling: Personal computer adoption and use as an illustration. *Technology Studies, 2*(2), 285–309.

Bar-On, R. (1997). *Bar-On Emotional Quotient Inventory: Technical manual*. Toronto: Multi-Health Systems.

Bar-On, R. (2004). The Bar-On Emotional Quotient Inventory (EQ-i): Rationale, description and summary of psychometric properties. In G. Geher (Ed.), *Measuring emotional intelligence: Common ground and controversy* (pp. 111–142). Hauppauge: Nova.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI) 1. *Psicothema, 18*(Suplemento), 13–25.

Boyatzis, R. E. (1982). *The competent manager: A model for effective performance*. New York: Wiley.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks: Sage.

Boyatzis, R. E. (2006). Using tipping points of emotional intelligence and cognitive competencies to predict financial performance of leaders. *Psicothema, 18*, 124–131.

Boyatzis, R. E. (2009). A behavioral approach to emotional intelligence. *Journal of Management Development, 28*(9), 749–770.

Boyatzis, R. E., & Gaskin, J. (2010). *A technical note on ESCI and ESCI-U: Factor structure, reliability, convergent and discriminant validity using EFA and CFA*. Boston: The Hay Group.

Boyatzis, R. E., & Goleman, D. (1996/1999). *Emotional competency inventory.* Boston: The Hay Group.

Boyatzis, R. E., & Ratti, F. (2009). Emotional, social and cognitive intelligence competencies distinguishing effective Italian managers and leaders in a private company and cooperatives. *Journal of Management Development, 28*(9), 821–838.

Boyatzis, R. E., & Saatcioglu, A. (2008). A twenty year view of trying to develop emotional, social and cognitive intelligence competencies in graduate management education. *Journal of Management Development, 27*(3), 92–108.

Boyatzis, R. E., Stubbs, E. C., & Taylor, S. N. (2002). Learning cognitive and emotional intelligence competencies through graduate management education. *Academy of Management Journal on Learning and Education, 1*(2), 150–162.

Boyatzis, R. E., Lingham, A., & Passarelli, A. (2010). Inspiring the development of emotional, social, and cognitive intelligence competencies in managers. In M. Rothstein & R. Burke (Eds.), *Self-management and leadership development* (pp. 62–90). Cheltenham, UK: Edward Elgar Publishers.

Boyatzis, R. E., Brizz, T., & Godwin, L. (2011). The effect of religious leaders' emotional and social competencies on improving parish vibrancy. *Journal of Leadership and Organizational Studies, 18*(2), 192–206.

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29*(9), 1147–1158.

Bray, D. W., Campbell, R. J., & Grant, D. L. (1974). *Formative years in business: A long term AT&T study of managerial lives*. New York: Wiley.

Campbell, J. P., Dunnette, M. D., Lawler, E. E. I. I. I., & Weick, K. E., Jr. (1970). *Managerial behavior, performance, and effectiveness*. New York: McGraw-Hill.

Camuffo, A., Gerli, F., Borgo, S., & Somia, T. (2009). The effects of management education on careers and compensation: A competency-based study of an Italian MBA programme. *Journal of Managemetn Development, 28*(9), 839–858.

Camuffo, A., Gerli, F., & Gubitta, P. (2012). Competencies matter: Modeling effective entrepreneurship in north east of Italy small firms. *Cross Cultural Management, 19*(1), 48–66.

Cavallo, K., & Brienza, D. (2002). *Emotional competence and leadership excellence at Johnson & Johnson: The emotional intelligence and leadership study*. Paper downloaded on 2 May 2002, from http://www.eiconsortium.org/

Cherniss, C. (2010). Emotional intelligence: Toward clarification of a concept. *Industrial and Organizational Psychology, 3*(2), 110–126.

Cherniss, C., & Boyatzis, R. E. (2013). Using a multilevel theory of performance based on emotional intelligence to conceptualize and develop "soft" leader skills. In R. Riggio & S. Tan (ed.). *Building Interpersonal Skills in Management Programs*. 53–72.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research, 14*(2), 189–217.

Ciarrochi, J. V., Caputi, P., & Mayer, J. D. (2003). The distinctiveness and utility of a measure of trait emotional awareness. *Personality and Individual Difference, 34*, 1477–1490.

Conte, J. M. (2005). A review and critique of emotional intelligence measures. *Journal of Organizational Behavior, 26*(4), 433–440.

Conte, J. M., & Dean, M. A. (2006). Can emotional intelligence be measured? In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 59–77). Mahwah: Lawrence Erlbaum.

Dawda, D., & Hart, S. D. (2000). Assessing emotional intelligence: Reliability and validity of the Bar-On Emotional Quotient Inventory (EQ-i) in university students. *Personality and Individual Differences, 28*, 797–812.

Dreyfus, C. (2008). Identifying competencies that predict effectiveness of R and D managers. *Journal of Management Development, 27*(1), 76–91.

Dulewicz, V., & Higgs, M. J. (1999). Emotional intelligence: A review and evaluation study. *Journal of Managerial Psychology, 15*(4), 341–368.

Dulewicz, V., Higgs, M., & Slaski, M. (2003). Measuring emotional intelligence: Content, construct and criterion-related validity. *Journal of Managerial Psychology, 18*(5), 405–420.

Fernández-Berrocal, P., & Extremera, N. (2006). Emotional intelligence: A theoretical and empirical review of its first 15 years of history. *Psychotema, 18*, 7–12.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327–335.

Fornell, C. A. (1982). *Second generation of multivariate analyses: Measurement and evaluation methods*. New York: Praeger.

Furnham, A., & Petrides, K. V. (2003). Trait emotional intelligence and happiness. *Social Behavior and Personality: An International Journal, 31*(8), 815–823.

Gignac, G. E. (2005). Evaluating the MSCEIT V2. 0 via CFA: Comment on Mayer et al. (2003). *Emotion, 5*(2), 233–235.

Gignac, G. E., Palmer, B. R., Manocha, R., & Stough, C. (2005). An examination of the factor structure of the Schutte Self-Report Emotional Intelligence (SSREI) scale via confirmatory factor analysis. *Personality and Individual Differences, 39*(6), 1029–1042.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Goleman, D. (1998). *Working with emotional intelligence*. New York: Bantam.

Grubb, W. L., III, & McDaniel, M. A. (2007). The fakability of Bar-On's emotional quotient inventory short form: Catch me if you can. *Human Performance, 20*(1), 43–59.

Gutierrez, B., Spencer, S. M., & Zhu, G. (2012). Thinking globally, leading locally: Chinese, Indian, and Western leadership. *Cross Cultural Management, 19*(1), 67–89.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1995). *Multivariate data analysis: With readings*. Upper Saddle River: Prentice-Hall, Inc.

Havers, G. (2010). *EI at the heart of performance: The implications of our 2010 ESCI research*. London: The Hay Group.

Hay Group. (2011). *Emotional and social competency inventory: A user guide for accredited practitioners*. Boston: The Hay group.

Higgs, M. J., & Dulewicz, V. (2002). *Making sense of emotional intelligence* (2nd ed.). Windsor: NFER-Nelson.

Hopkins, M., & Bilimoria, D. (2008). Social and emotional competencies predicting success for female and male executives. *Journal of Management Development, 12*(1), 13–35.

Howard, A., & Bray, D. (1988). *Managerial lives in transition: Advancing age and changing times*. New York: Guilford Press.

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*(1), 54–78.

Koman, L., & Wolff, S. (2008). Emotional intelligence competencies in the team and team leader. *Journal of Management Development, 12*(1), 56–75.

Kotter, J. P. (1982). *The general managers*. New York: Free Press.

Law, K. S., Wong, C., & Song, L. J. (2004). The construct and criterion validity of emotional intelligence and its potential utility in management research. *Journal of Applied Psychology, 87*(3), 483–496.

Leeper, R. W. (1948). A motivational theory of emotion to replace emotion as disorganized response. *Psychological Review, 55*(1), 5–21.

Lim, K., Sia, C., Lee, M., & Benbasat, I. (2006). Do I trust you online, and if so, will I buy? An empirical study of two trust-building strategies. *Journal of Management Information Systems, 23*(2), 233–266.

Lisicki, J. M., Jr. (2011). *An examination of the relationship between emotional intelligence and work related outcomes*. An unpublished doctoral dissertation, Capella University.

Lloyd, M. (2001). *Emotional intelligence and Bass Brewers Ltd*. Dissertation, Nottingham Business School.

Luthans, F., Hodgetts, R. M., & Rosenkrantz, S. A. (1988). *Real managers*. Cambridge, MA: Ballinger Press.

MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion, 8*(4), 540–551.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence. Science and myth*. Cambridge, MA: The MIT Press.

Matthews, G., Zeidner, M., & Roberts, R. D. (2004). *Emotional intelligence: Science and myth*. Cambridge, MA: MIT Press.

Matthews, G., Emo, A. K., Funke, G., Zeidner, M., Roberts, R. D., Costa, P. T., Jr., & Schulze, R. (2006). Emotional intelligence, personality, and task-induced stress. *Journal of Experimental Psychology: Applied, 12*(2), 96–107.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence* (pp. 3–34). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 2*, 267–298.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 3*, 97–105.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology, 59*, 507–536.

McClelland, D. C. (1973). Testing for competence rather than intelligence. *American Psychologist, 28*(1), 1–40.

McClelland, D. C. (1985). *Human motivation*. Glenview: Scott Foresman.

McClelland, D. C. (1998). Identifying competencies with behavioral event interviews. *Psychological Science, 9*, 331–339.

Mikolajczak, M., Luminet, O., & Menil, C. (2006). Predicting resistance to stress: Incremental validity of trait emotional intelligence over alexithymia and optimism. *Psicothema, 18*(Suplemento), 79–88.

Mikolajczak, M., Luminet, O., Leroy, C., & Roy, E. (2007). Psychometric properties of the Trait Emotional Intelligence Questionnaire: Factor structure, reliability, construct, and incremental validity in a French-speaking population. *Journal of Personality Assessment, 88*(3), 338–353.

Murensky, C. L. (2000). *The relationship between emotional intelligence, personality, critical thinking ability, and organizational leadership performance at upper levels of management*. Dissertation, George Mason University.

Murphy, K. R. (2006). Four conclusions about emotional intelligence. In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 345–354). Mahwah: Lawrence Erlbaum.

Nel, H. (2001). *An industrial psychological investigation into the relationship between emotional intelligence and performance in the call centre environment*. Unpublished master's thesis, University of Stellenbosch, Department of Industrial Psychology.

Nunnally, J., Bernstein, I., & Berge, J. (1994). *Psychometric theory*. New York: McGraw-Hill.

Palmer, B. R., Gignac, G., Manocha, R., & Stough, C. (2005). A psychometric evaluation of the Mayer-Salovey-Caruso emotional intelligence test version 2.0. *Intelligence, 33*(3), 285–305.

Petrides, K. V., & Furnham, A. (2000). On the dimensional structure of emotional intelligence. *Personality and Individual Differences, 29*, 313–320.

Petrides, K. V., & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality, 17*, 425–448.

Petrides, K. V., & Furnham, A. (2003). Trait emotional intelligence: Behavioral validation in two studies of emotion recognition and reactivity to mood induction. *European Journal of Personality, 17*, 39–75.

Petrides, K. V., Frederickson, N., & Furnham, A. (2004). The role of trait emotional intelligence in academic performance and deviant behavior at school. *Personality and Individual Differences, 36*, 277–293.

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology, 98*(2), 273–289.

Piel, M. A. (2008). *Emotional intelligence and critical thinking relationships to transformational leadership*. Unpublished doctoral dissertation, University of Phoenix.

Pittenger, L. (2012). *IT professionals: Understanding what impacts their desire to engage in work.* PhD, Case Western Reserve University, Cleveland.

Quinn, J. (2013). *The importance of compassion and vision in physician leadership.* PhD, Case Western Reserve University, Cleveland.

Ramo, L., Saris, W., & Boyatzis, R. E. (2009). The impact of emotional and social competencies on effectiveness of Spanish executives. *Journal of Management Development, 28*(9), 771–793.

Roberts, R. D., Schulze, R., O'Brien, K., MacCann, C., Reid, J., & Maul, A. (2006). Exploring the validity of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) with established emotions measures. *Emotion, 6*, 663–669.

Rossen, E., & Kranzler, J. H. (2009). Incremental validity of the Mayer–Salovey–Caruso Emotional Intelligence Test Version 2.0 (MSCEIT) after controlling for personality and intelligence. *Journal of Research in Personality, 43*(1), 60–65.

Rossen, E., Kranzler, J. H., & Algina, J. (2008). Confirmatory factor analysis of the Mayer–Salovey–Caruso Emotional Intelligence Test V 2.0 (MSCEIT).

Personality and Individual Differences, 44(5), 1258–1269.

Ryan, G., Emmerling, R. J., & Spencer, L. M. (2009). Distinguishing high performing European executives: The role of emotional, social and cognitive competencies. *Journal of Management Development, 28*(9), 859–875.

Ryan, G., Spencer, L. M., & Bernhard, U. (2012). Development and validation of a customized competency-based questionnaire: Linking social, emotional, and cognitive competencies to business unit profitability. *Cross Cultural Management, 19*(1), 90–103.

Saklofske, D. H., Austin, E. J., & Minski, P. S. (2003). Factor structure and validity of a trait emotional intelligence measure. *Personality and Individual Differences, 34*, 707–721.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185–211.

Salovey, P., & Mayer, J. D. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications*. New York: Basic Books.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167–177.

Sergio, R. P. (2001). *Emotional intelligence and mental ability as determinants of job performance among plant supervisors in selected manufacturing firms*. Unpublished master's thesis, De La Salle University-Dasmariñas, Behavioral Sciences Department, Philippines.

Sevinc, L. (2001). *The effect of emotional intelligence on career success: Research on the 1990 graduates of business administration faculty of Istanbul University*. Unpublished master thesis, Istanbul University.

Shams, M. (2008). *The impact of the leaders' emotional intelligence on the employee's job performance and job satisfaction*. Unpublished bachelor dissertation, German University, Cairo.

Sharma, R. (2012). Measuring social and emotional intelligence competencies in the Indian context. *Cross Cultural Management, 19*(1), 30–47.

Spector, P. E., & Johnson, H. M. (2006). Improving the definition, measurement, and application of emotional intelligence. In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 325–344). Mahwah: Lawrence Erlbaum.

Spencer, L. M., Jr., & Spencer, S. M. (1993). *Competence at work: Models for superior performance*. New York: Wiley.

Stagg, G., & Gunter, D. (2002). *Emotional intelligence in the fire service*. Unpublished Paper, London Fire Brigade.

Sternberg, R. J. (1996). *Successful intelligence: How practical and creative intelligence determine success in life*. New York: Simon and Shuster.

Thornton, G. C., III, & Byham, W. C. (1982). *Assessment centers and managerial performance*. New York: Academic.

Victoroff, K., & Boyatzis, R. E. (2013). An examination of the relationship between emotional intelligence and dental student clinical performance. *Journal of Dental Education, 77*(4), 416–426.

Williams, H. (2008). Characteristics that distinguish outstanding urban principals. *Journal of Management Development, 27*(1), 36–54.

Wolff, S. B. (2007). *Emotional and social competency inventory: Technical manual up-dated ESCI research titles and abstracts*. Boston: The Hay Group.

Wong, C. S., & Law, K. S. (2002). The effect of leader and follower emotional intelligence on performance and attitude: An exploratory study. *Leadership Quarterly, 13*, 243–274.

Young, M., & Dulewicz, V. (2009). A study into leadership and management competencies predictive superior performance in the British Royal Navy. *Journal of Management Development, 28*(9), 794–820.

Intelligence as a Malleable Construct

18

Lisa S. Blackwell, Sylvia Rodriguez, and Belén Guerra-Carrillo

Introduction

The academic debate about the nature of intelligence resonates broadly not only with educational practitioners and policy makers but also with the general public. Most people think they know what intelligence is and that they know it when they see it. But what do people mean when they talk about being smart, brilliant, or clever? And why does it matter so much?

In this chapter, we focus on research showing how the way that people think about intellectual ability drives the choices they make and the outcomes they achieve, sometimes in counterintuitive ways. We show how a person's concept of intelligence can impact both their performance on cognitive tasks in the short run and their achievement over the longer term, and why this is so. We review the evidence from cognitive neuroscience for these different conceptions of intelligence. Finally, we discuss how such concepts can be influenced and changed and the practical implications of this research for educational policy and practice.

Why do people care so much about the nature of intelligence? Traditionally, particularly in western cultures, intelligence has been seen as the golden ticket to success. If you had a good amount of it, you would be rewarded with educational, professional, and financial success, and those with a great deal—the geniuses among us—would attain eminence and make a mark on posterity. Implicit in this view is the idea that intelligence is a "gift"—an innate attribute that one possesses in a relatively fixed quantity, for better or worse. Historically, the relatively high stability in individual performance on intellectual assessments over time and across tasks has led many to assume that this view of intelligence as fixed is correct (see, e.g., Bartels et al. 2002; Canivez and Watkins 1998; Herrnstein and Murray 1994; Hertzog and Schaie 1986), despite the strong dissent of original developers of the first IQ test, Alfred Binet and Theodore Simon (Binet 1975; Wolf 1973).

Unfortunately, in this case, believing may make it so. Because our society has presumed that intellectual gifts are innate and could be measured accurately, our education system has traditionally been structured to identify those students with apparent above-average intelligence, enrich their instruction, and track them into ever-greater opportunities, while those with presumed below-average ability were channeled into programs that

L.S. Blackwell, Ph.D. (✉) • S. Rodriguez, Ph.D.
Mindset Works, Inc., 340 S. Lemon Ave. #6463, Walnut, CA 91789, USA
e-mail: lasblackwell@mindsetworks.com

B. Guerra-Carrillo
132 Barker Hall, University of California at Berkeley, Berkeley, CA 94702, USA
e-mail: belyguerra@berkeley.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_18,

would prepare them for lower-skilled jobs (Borland 2003, 2005; Borland and Wright 2001; Darling-Hammond 1994, 1995; Kaufman 2013; Nisbett 2009). The result is that those who score well on performance measures early on are in fact generally offered more opportunities to cultivate their intellectual ability than those who do not and, at least partly as a result, often do become more skilled and successful, reinforcing our common paradigm of innate ability.

It turns out that something similar happens in the psychology of the individual as well. Our research, and that of many colleagues, shows that people's "theory" of intelligence—whether they believe it to be fixed or a malleable quality—influences the learning opportunities they will pursue, the effort they will invest, and their resulting growth. It can even impact how their brains function.

Mindsets About Intelligence

In a ***fixed mindset*** (often referred to as an ***entity theory*** in the research literature), people believe that their intelligence is relatively fixed, and there is not much that they can do to develop it. They agree with statements such as "you can learn new things, but you cannot really change your basic amount of intelligence." In contrast, in a ***growth mindset*** (also known as an ***incremental theory***), people believe that their intelligence is something they can change and develop incrementally over time. They agree with statements such as "you can always greatly change how intelligent you are" (Dweck 1999). As we will show, these different mindsets about intelligence drive the goals that people hold, the challenges they will tackle, the effort they will expend, their persistence in the face of difficulty, and, as a result, their performance and achievement over time (Blackwell et al. 2007; Dweck 1999; Dweck and Leggett 1988; Henderson and Dweck 1990).

Mindsets and Motivation

Over the past few decades, a wealth of research shows that, even when people demonstrate equal intellectual ability and skill, their beliefs about intelligence shape their responses to intellectual challenge. For those who hold a fixed mindset, the conception of intelligence as a fixed, uncontrollable quantity (of which they may have a lot or a little) orients them toward measuring and obtaining a positive evaluation of their ability. Thus, their primary goal is usually to perform well in order to appear smart—or at least to avoid performing poorly and looking dumb (Blackwell et al. 2007; Dweck and Leggett 1988). They tend to think that things come easily if one is smart and that effort is both a sign of low ability and relatively ineffective in overcoming it (e.g., Blackwell et al. 2007; Hong et al. 1999). When they experience a setback or failure, they are likely to attribute it to low ability rather than effort (e.g., Henderson and Dweck 1990), doubt their ability to recover, and manifest a "helpless" response, withdrawing effort and giving up rather than risking further exposure as unintelligent or untalented (e.g., Robins and Pals 2002).

On the other hand, those who hold a growth mindset, in which intelligence is a malleable quality that can be cultivated, are more focused on learning (thus increasing their ability) as a goal, even if it requires effort, struggle, and errors along the way (Dweck 1999; Dweck and Leggett 1988). They consider effort to be a pathway to development (e.g., Hong et al. 1999), and when they experience setbacks, they attribute them to lack of sufficient effort and in turn adopt a mastery-oriented approach, increasing their effort and taking on new study strategies (e.g., Robins and Pals 2002).

Thus, the different mindsets about intelligence set up different frameworks or "meaning systems" (Hong et al. 1999) for interpreting situations that involve learning, effort, challenge, and evaluation. Furthermore, it is when making a transition to a situation that poses ongoing, increasing challenge (where success is more difficult and less certain) that these mindsets have the greatest impact on behavior and achievement.

In a comprehensive longitudinal study with urban, largely minority students, we examined how students' mindsets set up contrasting motivational frameworks and academic outcomes as they made their way through a challenging transition to junior high school (Blackwell et al. 2007).

We studied three waves of students over three successive years, assessing their mindsets at the beginning of their seventh grade year and then following each wave as they made their way through the following two years of school. First, we examined how their mindsets were related to their goals in school, their attitudes toward effort, and their responses to failure. Analyses showed that, as found in prior studies, students with a growth mindset had stronger learning goals than the fixed mindset students—for example, they said that "It's much more important for me to learn things in my classes than it is to get the best grades"—and had much more positive attitudes toward effort, agreeing that "when something is hard, it just makes me want to work more on it, not less." Students with a fixed mindset, on the other hand, were more likely to say that "If you're not good at a subject, working hard won't make you good at it," and "When I work hard at something, it makes me feel like I'm not very smart."

How did these two groups of students feel about failure? These mindsets, goals, and beliefs about effort in turn predicted how students said they would respond to a poor grade on a quiz: the growth mindset students showed a clear mastery-oriented response, saying that they would "work harder in this class from now on" and "would spend more time studying for the next test." In contrast, many of the students with a fixed mindset had a helpless response—for example, saying they would "spend less time on this subject from now on," with some even admitting that they "would try to cheat on the next test" rather than risk another failure!

Mindsets and Achievement

How did these different mindset frameworks impact achievement over this challenging transition? Based on their prior sixth grade test scores, when they were in the less-challenging elementary school environment, the fixed and growth mindset students had similar levels of math skills upon entry into junior high school. But by the end of the first term, they began to pull apart, with the growth mindset students performing better, and these diverging trajectories continued over the next two years, widening the gap between the two groups each term (Fig. 18.1).

We examined the pathway from mindset to achievement outcomes using hierarchical linear modeling and found that the beliefs, goals, and attitudes that led to different patterns of behavior were responsible for the diverging trajectories of grades. The increasing challenge level, particularly in the math curriculum of a health science-focused school, spurred the students with a growth mindset to focus on learning, work harder, and use positive strategies when they encountered difficulty, with the result that they mastered the curriculum better than those who entered with a fixed mindset, despite the fact that both groups began with similar skills (Fig. 18.2).

What do these mindsets sound like in the words of real students? A rising eighth grader with a growth mindset explained how he thought about intelligence as a product of one's choices and behaviors, inextricably tied to learning and effort:

> Well, you can change it [your intelligence] because people are different. One year they can be lazy in school and the other year they're like, "All right, I've got to step it up because I want to get into college." … What makes me feel smart is participating and doing my homework and everything, because then I know that I'm doing my best.

Asked whether he liked schoolwork that made him think hard, he emphasized the value of challenge to his growth:

> Yes, I do, because it gives me a challenge and also it'll help me a lot and I can do better with it and everything.

Contemplating the prospect of failing a test, he immediately began seeking solutions based on effort:

> I would feel really bad, but at the same time I wouldn't be surprised because maybe the year before that you did really good, and then you know like you just put that same amount of effort. But like that year, the new year, it gets harder and everything… Maybe there was some notes that you could write down but you didn't bother because you already knew them. Maybe you didn't have it all memorized, so you forget some of the stuff. I guess what I would do was maybe work harder,

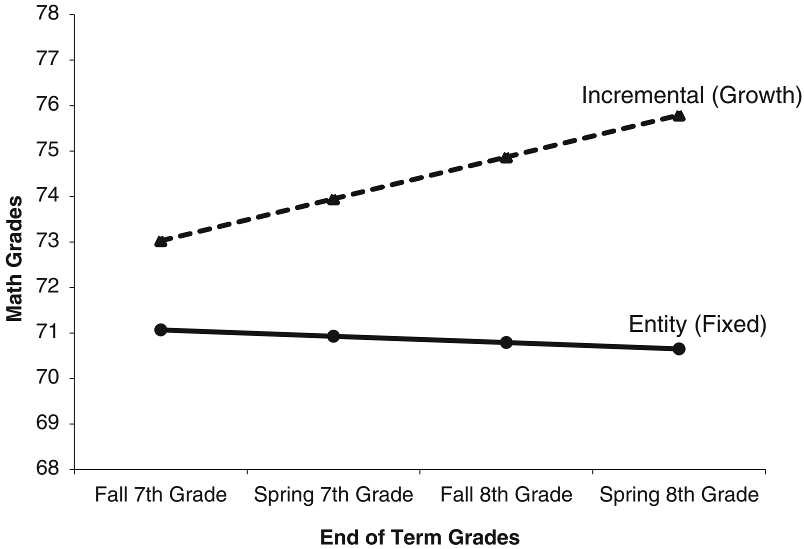


Fig. 18.1 Trajectory of middle school achievement as a function of student mindset in Study 1 (Notes: Growth and fixed mindset groups entered seventh grade with equal math achievement based on sixth grade test scores (not shown). They began to diverge by the end of the fall term of seventh grade (mid-year). By the end of eighth grade, the achievement gap was 10 % of the total grade points that differentiate a failing grade (60 %) from a perfect score (100 %). Reprinted from Blackwell et al. 2007, p. 251. Reprinted with permission. Copyright 2007 from Society for Research in Child Development, Inc.)

> and start thinking, Oh, wow. Okay, so I didn't do so good but maybe this time I can do good.

In contrast, his classmate laboring under a fixed mindset talked about her uncertainty about her ability to learn and how it made her feel helpless:

> Well I'm going to have to probably agree [that you can't change your intelligence] because sometimes– well for me there's limits on what I can learn and what I can't… I tend to space out a lot. And when I space out it's like the teacher will ask me a question and I have no idea what she's saying. And so I just have to sit in silence until she gives up and picks somebody else.

When asked what made her feel smart, she looked to external validation through getting the "right answer" and admitted that she preferred things she could do easily versus challenging work:

> Like say I got a question right in front of the whole class, then that makes me feel like kind of smart and special … I think it's so much easier and quicker if you know it by heart and you just do it right away and get it over with … Over-thinking sometimes can just really frustrate me. What I've done is I would just give up and my friends would sometimes give the answers to me.

When contemplating failure, she shared a recent incident and her collapse in the face of challenge:

> I was doing my test and what happened is I was reading this question that I really didn't know … from there on I just circled randomly and I just completely gave up on them, even like trying on the test.

The motivational implications of these two different frameworks, and their resulting impact on performance and achievement, have been demonstrated in many studies spanning kindergarten through graduate school (Aronson et al. 2002; Blackwell et al. 2007; Dweck and Leggett 1998; Good et al. 2003; Heyman et al. 2003; Kray and Haselhuhn 2007; Smiley and Dweck 1994; Yeager et al. 2013). Over and over again, researchers have shown that the way people think about their intelligence can become a self-fulfilling prophecy, expanding or limiting their motivation, growth, achievement, and, ultimately, their ability.

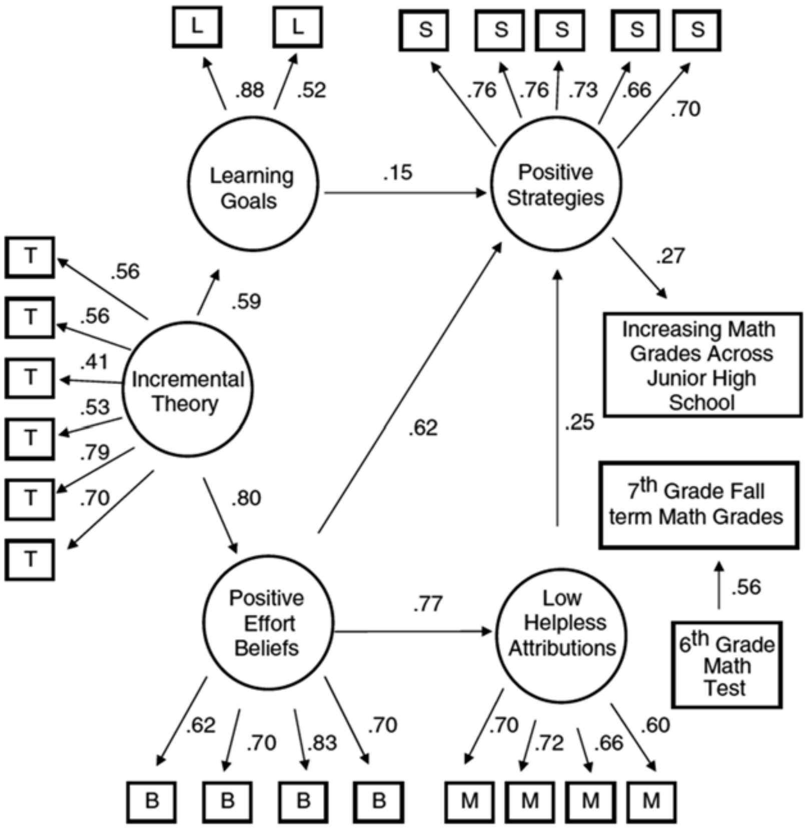


Fig. 18.2 Process model depicting the relations between student mindset, other beliefs and behaviors, and achievement in Study 1 (Notes: The more firmly students held a growth mindset (incremental theory), the more they endorsed learning goals and positive beliefs about effort. These goals and beliefs were associated with positive learning strategies and resilient responses to challenge, which in turn predicted greater math achievement gains. Reprinted from Blackwell et al. 2007, p. 253. Reprinted with permission. Copyright 2007 from Society for Research in Child Development, Inc.)

How Malleable Is Intelligence Really?

Without doubt, people can gain knowledge and skills through learning, but can they really develop their intelligence as we understand it? A robust debate about the true nature of intelligence continues (see, e.g., Kaufman 2013; Nisbett 2009), but most people think of intelligence as a generalized capacity for learning and reasoning that can be assessed by instruments such as IQ tests. Without weighing in on that complex question, we can agree that the version of intelligence measured by standard IQ tests is the result of combining scores from various subtests that measure a wide variety of knowledge and cognitive processes that are highly intercorrelated, such that if you score well on one, chances are that you will also score well on another. James Flynn (2007) explained this calculation with a clever analogy comparing it to measuring performance in a decathlon, where performance is computed from 10 events that each assess a different ability. For example, strength can be calculated from performance on throwing events, while speed can be assessed through sprinting events. Similarly, different subtests of intelligence assessments measure cognitive factors such as our ability to maintain and manipulate information in mind (working memory), integrate features of and consider relationships between stimuli (reasoning), and process information fluidly (processing speed), among others. A portion of these subtests may also measure the accumulation of knowledge about the meaning of words or arithmetic rules (Naglieri and Goldstein 2009). Intuitively, one can suspect that what we do or are exposed to in our daily life could influence how well we score on one or many of these different subtests and subsequently affect how *intelligent* we are deemed to be. However, for a long time, it was believed that intelligence was something we inherited and could not do much to change (see, e.g., Herrnstein and Murray 1994).

The evidence for the primacy of innate ability has not been well supported by accounts measuring population changes in IQ performance since the inception of the Weschsler Intelligence Scale, one of the main measures of intelligence. The well-documented Flynn effect (Flynn 2007) describes how IQ scores on multiple well-established assessments of intelligence have been on the rise—in some instances, dramatically—from generation to generation, even on assessments that are deemed to be largely "culture-free." A compelling interpretation is that the performance capacities measured by these tests function as skills that can be improved and shaped by experience and schooling and that these experiences have shifted over time in a way that has changed how and what is learned by the majority of the population (Flynn 2007; Nisbett 2009).

In fact, over the past century, various studies conducted all over the world have documented the role that schooling plays in cultivating students' intelligence. If intelligence is a fixed ability, environmental experiences, such as educational enrichment, should not alter it. Yet, countless examples confirm the finding that, relative to children who remain in school, those who are denied educational experiences often display a gradual but persistent decline in performance on intelligence measures—as much as 6 IQ point decrements for every year of schooling lost (see Ceci 1991; Nisbett 2009 for a review). Similarly, related environmental factors such as socioeconomic status have been found to predict individual change in IQ, with low-SES children showing a decrease in IQ over time (Breslau et al. 2001).

To what extent are scores on assessments of these sorts malleable, and how does experience and learning impact performance on them? Over the last few years, research aimed at answering this question has provided strong evidence suggesting that cognitive skills such as those tested by intelligence tests can in fact improve with practice.

In one study, elementary school children at a low SES school played with one of two sets of board games and video games for 8 weeks (Mackey et al. 2011). In the first group, children played with games that engaged their reasoning ability, such as games that asked them to consider and integrate multiple rules or complete patterns of shapes. A second group of children played with games that involved processing speed, which required them to make motor responses to visual cues as fast as possible following simple game rules. At the beginning and again at the end of the 8 weeks, tests in relevant areas pertaining to either reasoning ability or processing speed were administered. The results were striking. Children who had played with the reasoning games increased their nonverbal reasoning by 32 %, which translates roughly to a 10-point increase in IQ. On the other hand, those who had played with the processing speed-focused games increased their processing speed by 27 %, but not their reasoning skills. These results demonstrate not only that IQ scores can change significantly in a short period of time but, more importantly, that targeted interventions can improve performance on the cognitive processes assessed in intelligence measures, contradicting the belief that our intelligence is fixed.

Studies examining other cognitive functions measured by intelligence tests have shown converging evidence that speaks to the malleability of these constructs. For example, Brehmer et al. (2012) asked a group of adults to use an adaptive computerized program to train working memory and compared them to a control group that used a nonadaptive, low difficulty working memory training. Before and after the training period, participants completed assessments of working

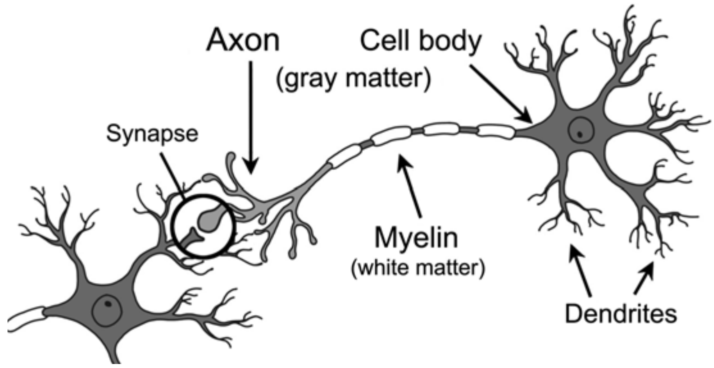


Fig. 18.3 Illustration of neuron with connections (Copyright 2013 from Mindset Works, Inc.)

memory, which were the same as those administered as part of the Weschsler Intelligence Scale. As part of the training, participants practiced maintaining in mind over short periods of time multiple bits of information (words or locations of objects in space). They also sequenced these items in a particular order according to the exercise's instructions. For those in the adaptive training program, the quantity of information that needed to be held or manipulated in working memory changed depending on their performance, increasing (and therefore becoming more difficult) as participants became more proficient. After five weeks of training, participants in this adaptive training group showed significant improvements in their working memory performance. These improvements were greater compared to the group that did not receive an adaptive training, even though both groups began their training with similar scores. Interestingly, in a follow-up assessment, researchers found that the performance gains made by the adaptive training group were maintained three months after the training period. These findings are especially impressive given the brief nature of the intervention.

In sum, these examples support the idea that the cognitive skills measured in intelligence tests can be improved and maintained by targeted training using increasing demands (challenge). The significance of the flexibility of these cognitive functions extends beyond improvements in a test score—they speak to our capacity to continue to adapt to new cognitive demands imposed by our environment.

Neuroplasticity and Brain Function

How do learning-based gains in cognitive performance relate to the modifiability of neural structures in the brain? Research in the neuroscience of learning demonstrates significant plasticity in brain structure and function and shows that these changes are highly dependent on both behavior and environment.

The brain is composed of billions of specialized cells known as neurons (see Fig. 18.3). Neurons form part of the body's communication hub, processing, sending, and receiving vital information through an interconnected network. Surrounding a neuron's center are a series of extending branches known as dendrites that oversee and transport the collection of chemical messages received from other neurons. Collected information is eventually converted into an electrical impulse that travels down a long channel known as the neuron's axon. This axon branches out into smaller axon channels that each culminate in a small bulb that forms part of a structure called a synapse. It is at a synapse where the impulse triggers the release of chemical substances, or neurotransmitters, from the neuron. These neurotransmitters then enter the receiving neuron through receptors of the receiving neuron.

Although brain plasticity may not be specific to neural changes, much of the research has focused on how the connectivity between neurons can change with experience. Studies on animals have informed us about the changes that occur at a cellular level that may modify the connection of neurons within and between brain regions. Based on those findings, Zatorre et al. (2012) discuss some of the changes that may happen at the cellular level and underlie plasticity in the human brain. For example, axons can become insulated with myelin, a fatty tissue that makes up the "white matter" of the brain and speeds the transmission of electrical impulses, or their existing myelin covering can become thicker. Also, groups of axons projecting between brain regions can become more organized, thus improving their connectivity. In humans, one or many of these transformations may reflect changes in the integrity of the microstructure of the brain's white matter, which contains groups of myelinated axons whose cell bodies form the brain's gray matter. In addition to speeding up their transmission, neurons can improve their connectivity by increasing their surface area of connections through the creation of new neurons and synapses or branching of dendrites. These events may be reflected in the structural and functional changes of the gray matter after repeated practice of a physical or mental skill.

An example of our brain's ability to adapt to slight modifications to our daily activities can be seen in a seminal study where a group of individuals with no juggling experience were taught how to juggle over a period of three months. In comparison to their own brain scans taken at the beginning of the study and also to those of a group of individuals who were not taught how to juggle, the post-training brain scans of the jugglers showed an increase in the gray matter thickness of brain regions that support the ability to perceive motion and anticipate where objects will be in space (Draganski et al. 2004).

Changes in the brain are not limited to developing a new visual-motor skill. Exciting findings from studies where participants practiced a cognitive process have revealed the brain's ability to adapt to different cognitive demands. Using functional magnetic resonance imaging (fMRI), which measures changes in blood oxygenation and flow in the brain associated with neural activity, Mackey et al. (2013) measured the neural effects of intensive reasoning training in young adults. After a three-month law school admission exam (LSAT) preparation course, in which a little over 60 h were devoted to practicing problems that relied heavily on reasoning, the fMRI scans showed changes in the intrinsic connectivity of the student's brains. The measure of intrinsic connectivity, known as resting-state fMRI, is thought to reflect repeated history of synchronized activity between regions, since the scan is captured at resting state, when a person is asked not to engage in any task. The group of individuals who underwent the reasoning training showed greater intrinsic connectivity between areas involved in reasoning skills compared to their well-matched controls, and a specific pattern of these connections was related to greater improvements in their LSAT scores.

In addition, these researchers also examined the changes in white matter microstructure resulting from this training program (Mackey et al. 2012). They found an increase in the coherence of white matter tracts connecting regions of the brain that support reasoning skills, reflecting the integrity of the structure of white matter discussed earlier. Although the specific mechanisms behind these changes are unknown, they are, nonetheless, thought to reflect strengthening of the connectivity between brain regions brought on by experience or development.

The brain's malleability also makes it susceptible to negative factors, such as stress or unstable home environments (see, e.g., Erikson et al. 2003; Hackman and Farah 2009; Lupien et al. 2009). To counteract this, Neville and colleagues (2013) developed an 8-week intervention targeting selective attention, the ability to control where our focus is directed, in part aimed at increasing school readiness for preschoolers of low socioeconomic status. They reasoned that because a stressful environment and more inconsistent parenting practices are often more prevalent in low-SES compared to higher-SES households, training preschoolers' primary caregivers might also be beneficial. Thus, they compared three groups of children. One received their preschool education as usual. A second group received attention exercises only. A third group of children also received attention exercises, and their parents received training in a curriculum targeted to develop family stress regulation and other strategies aimed at improving the way parents interacted with, disciplined, and facilitated their children's attention. To explore the impact of the training on cognitive functioning, researchers used electroencephalography (EEG)—a noninvasive measure with excellent temporal sensitivity that can be used to capture changes in electrical potentials occurring within the brain that are elicited from the scalp. The signal embedded in this EEG, known as event-related potentials (ERPs), can map out attentional and conceptual processes that emerge in response to specific task-related events, such as hearing a particular sound. In a test where children were asked to focus their attention on only one of two stories played simultaneously, early attention ERP components related to probes embedded in the stories showed that children who received the family-based intervention were more successful at focusing their attention on the story they were instructed to attend. Additionally, this group also improved in measures of nonverbal IQ, while their parents showed lower levels of stress.

Together, these studies suggest that experience and learning can result in tangible, measurable impacts on the brain and in turn on a variety of cognitive functions. This is especially promising because these brain changes are seen in response to small changes in the experiences a person engages with, such as practicing a specific skill over a short period of time. Given these findings, we can anticipate that, as they continue to work and study, virtually all students can continue to develop their abilities over time through positive behaviors like effort and practice, in a way that would ultimately be evidenced by changes to brain structures and activity.

The evidence of performance improvement and neuroplasticity seen in these studies lends support to the concept of malleable intelligence that underlies a growth mindset. Further, the fact that such changes are the result of behaviors such as deliberate practice and engaging with increasingly difficult tasks helps illuminate why the increased effort, challenge seeking, and persistence associated with a growth mindset would result in higher achievement. Intelligence can be developed—but only if it is exercised. But our colleagues and we wondered whether engaging in overt behaviors, such as practicing and tackling more challenging tasks, were the only way that mindsets could impact learning—or whether the beliefs and goals that make up the different mindsets might directly influence the way the brain processes information.

The Neuroscience of Mindsets

What is happening in the brain when people are laboring under the different mindsets? Researchers have begun to explore some of the neural mechanisms underlying a growth and fixed mindset.

A fascinating series of studies looking at brain activity in relation to mindset and different performance conditions showed that mindsets can lead to different patterns of observed activity in the brain, with consequences for cognitive functioning. Moser et al. (2011), for example, tracked how students allocated their attention while completing a task that required continuous monitoring and responding to a target displayed on a computer screen. How did students with different mindsets react, especially after making a mistake? To explore this question, researchers looked at specific ERPs that have been previously mapped to attention and awareness to errors. They found that individuals with a growth mindset were more likely to attend to the errors they made than those with a fixed mindset and were also more likely to improve their accuracy on the next trial.

Interestingly, additional analyses revealed that their attentional response mediated their performance. In other words, it was *because* participants with a growth mindset oriented their attention to errors that they did better on the task. These findings showed that people with a growth mindset were more successful at reorienting their attention to the task at hand and suggest that they were not discouraged by errors but responded in an adaptive way that allowed them to persist and improve.

Individuals with a fixed mindset may also orient attention to errors when there is salient negative feedback, but may do so in a way that ultimately undermines learning (Mangels et al.

2006). An ERP study by Jennifer Mangels and her colleagues found that individuals with a fixed mindset showed an enhanced awareness of and orientation towards errors made on a challenging general knowledge question task, in which individuals received accuracy feedback (whether their response was correct or incorrect) followed by learning feedback (the correct answer) after each question. However, unlike the growth mindset group in the previous study, this orienting did not aid their performance. Fixed mindset participants showed a neural response to learning feedback that was indicative of lower success at encoding the correct answer or storing and committing the information to memory. In fact, in a surprise retest of all the items that they had previously answered incorrectly, fixed mindset participants corrected fewer items than their growth mindset counterparts.

These findings illustrate how people's mindsets may differentially impact their attentional response, particularly following challenge. On one hand, individuals harboring fixed mindset thinking may inadvertently set themselves up for failure by directing their attention to their performance and discounting an opportunity to learn from their mistakes, whereas this does not seem to be the case for those with a growth mindset.

As previously discussed, students with a fixed mindset often hold a performance focus in which they are particularly concerned with the goal of proving their abilities and achieving highly, especially in comparison to others, whereas those with a growth mindset typically endorse goals of learning and mastery (Blackwell et al. 2007; Dweck and Leggett 1988; Dweck 1999). What happens when students find themselves in academic contexts that promote either a performance or mastery goal? As students navigate from one academic context to the other, it is very possible that they may be receiving different messages from their environment about what is valued in each domain, potentially impacting how and what they learn.

In support of this possibility, recent ERP research finds that students do indeed have very different neural experiences when they encounter a mastery- versus performance-based context (Rodriguez et al. 2014). We recruited undergraduates to complete a challenging general knowledge task drawn from Mangels et al. (2006). The task, which contained two blocks of questions, prompted students to complete all questions in a block before being presented with the second block. Importantly, as students were presented with a block, they first read task instructions that differentially framed it. In the *performance* frame, students read instructions that oriented their focus on their accuracy and how their performance would be compared to that of other university students. However, in the *mastery* frame, these same students were instead asked to focus on those questions that they found most interesting and learned the best from rather than on their performance. How would students respond to these two different but comparably challenging situations?

When task instructions emphasized performance and proving one's ability relative to others, students completing the difficult task displayed a neural response to corrective information (i.e., the correct answer) following an error that was consistent with processes related to superficial encoding of that information. However, these very same students, when completing a task that instead emphasized learning and mastery, experienced a neural response to the correct answer (after an error) that was consistent with processes that reflect deeper encoding of that information.

This work suggests that although learning may occur in both mastery- and performance-based contexts, the nature of that learning may be very different. In a performance environment, students may attend to problem solving only insofar as it allows them to get the right answer, understanding it only at its surface, whereas in a learning environment, students may not be solely focused on their outcomes and instead orient their efforts to understanding the content in a manner that may ultimately contribute to longer-term retention. These findings are especially intriguing since these different neural processes, which are suggestive of qualitatively different kinds of learning, emerged within person after just brief exposures to each frame. Thus, they provide continued support for students' sensitivity and differential response to input from their learning environment. As we will see, messages conveyed to students through their interpersonal experiences with others can also powerfully shape their beliefs and behaviors in both the short and long term.

Mindsets and the Influence of Others

As we have seen, the mindsets that individuals carry with them affect their goals, cognitive functioning, motivation-related behavior, and academic outcomes. However, these mindsets themselves are not fixed. The messages that people get from others in their environment can influence their mindsets and impact motivation and performance in immediate, powerful, and often surprising ways. In fact, it turns out that the very messages that one might think would be most encouraging—such as praise for intelligence—can actually undermine performance on intellectual tasks.

A pioneering series of studies by Claudia Mueller and Carol Dweck (1998) examined the impact of praise on fifth graders' challenge seeking and performance. Mueller and Dweck had the children complete a set of puzzles drawn from Raven's Progressive Matrices (a common measure of nonverbal reasoning). Initially, they gave them problems matched to their grade level and children solved most of them successfully. Then the researchers praised the students for their performance. They told one group of randomly chosen children, "Wow, that's a really good score. You must be smart at this" (***intelligence praise*** condition). A second group was told, "Wow, that's a really good score. You must have worked hard at this" (***effort/process praise*** condition). (Process praise can refer to anything about the process the child engaged in: their strategy, focus, effort, choices, or perseverance.) Then they looked at how these different kinds of praise would affect the students' behavior and performance. First, they asked them which type of puzzle they would prefer to do next: an easy one, like the ones they had done, on which they would perform well, or more difficult ones, from which they would learn. While the children praised for process overwhelmingly chose the more difficult ones, the majority of children praised for intelligence chose to repeat the same easy puzzles! Rather than giving children the confidence to tackle a challenge, praise for intelligence had actually made them want to stay in their safe zone, even though it meant that they would learn nothing new.

Clearly, over the longer term, sacrificing such learning opportunities could have a negative impact on skill development. But strikingly, the praise also had an immediate effect on the children's intellectual performance. To test the impact of praise on the children's performance and resilience following challenge, the researchers next had the students work on more difficult puzzles, on which all the students struggled. They then gave them another easier set, similar to the first. How would they perform? The differences were telling. The students praised for effort improved significantly on the easy puzzles over their performance on the first trial (perhaps honing their skills on the more difficult problems). But those who had been praised for intelligence performed *worse* on the second attempt—they had lost confidence that they were smart at puzzles, and so they performed poorly. It is particularly notable that the Raven's task is one used to measure "fluid" intelligence (often considered to be an inherent problem-solving ability) and has often been used to assign children to gifted programs, yet it turned out that performance on this test could be undermined (or enhanced) by a single sentence. Intelligence praise activated a fixed mindset framework, along with the goals of looking smart and succeeding without effort, and produced a "helpless" response to challenge and an immediate decrease in apparent ability.

More recent studies have replicated and extended these findings. For example, one study explored how feedback linking success on an upcoming challenging activity to a group's supposed inherent ability impacted kids' performance on that task (Cimpian et al. 2012). The researchers found that when young (4–7-year-old) children were told that either girls or boys were really good at a game, the children, regardless of their gender, underperformed on the game, especially on more difficult items, relative to children who were provided with other kinds of instructions that did not suggest the inherent ability of groups. Cimpian and his colleagues argue that this occurred because children came to attribute their performance on the task as being innately linked to something out of their control—a "fixed" aspect of their identity, which in turn led to their underperformance.

Since the Mueller and Dweck (1998) studies, other researchers have investigated the impact of praise in real-world contexts on mindsets and performance over longer periods of time. They have found that parents can play a significant role in formulating their children's beliefs about their ability, sometimes with long-term effects. One longitudinal study found that children whose mothers gave them more process praise at 14–38 months were more likely to display a growth mindset and a greater desire for challenge at 7–8 years old (Gunderson et al. 2013). In a similar study with older children, 8–12-year-olds whose mothers praised them for ability were more likely to exhibit fixed mindset thinking and a reduced desire for challenge six months later (Pomerantz and Kemper 2013).

As the praise studies suggest, teachers too are in a position to influence their students' mindsets. For example, just as praise for intelligence can backfire, the way that we console children when they struggle may inadvertently trigger the fixed mindset pattern. Aneeta Rattan and her colleagues (Rattan et al. 2012) asked adults (some of whom were math teachers) to imagine a student in their class who had gotten a poor grade on the first math test of the year and to report how they would respond to the student. The adults in a fixed mindset were significantly more likely to try to console the student by saying that not everyone could be good in math—a message that students reported would lead them to conclude that they have low ability and to feel like giving up. However, adults in a growth mindset were more likely to urge students to try harder, and they gave them practical recommendations for strategies to achieve mastery.

In these ways, the mindsets that adults hold can be a factor in students' success. Indeed, a study looking at the impact of teacher mindsets on student achievement found that when teachers had a fixed mindset, students who had entered their class as low achievers remained so. In contrast, when teachers had a growth mindset, many of the students who had started the year as low achievers showed remarkable progress (Rheinberg et al. 2000). Thus, for the fixed mindset teachers, their experience confirmed their beliefs, as their students' relative status remained unchanged, whereas the growth mindset teachers saw their confidence in students' ability to grow realized. Once again, the mindsets that people hold can become self-reinforcing.

Crucially, a growth mindset is not the same as self-confidence or drive to achieve. In fact, as we saw in the praise studies, successful students who derive much of their self-esteem from performing well can be vulnerable when they encounter challenge or difficulty—particularly if they are in an evaluative context. Here, a young student about to begin middle school seems to have a robust sense of self-confidence as he explains his relish for a challenge—but his primary goal is to demonstrate his ability, rather than to learn:

> Right now my favorite subject is math because I'm really good at it – my grades in my report card have been really high in math so that's why I like it … A lot of times kids don't really want to work hard. But if you really want to know something new and really get good at it, you're going to have to work hard. And some kids like working hard. So it's like a challenge, like a puzzle… I like doing stuff that I'm good at and I know I'll get a good grade. And then I like thinking because I like – I'm the type of guy that likes puzzles and I like challenges 'cause it makes you feel like you're up to it and you're showing the person what you can do and it makes you feel good.

This student enjoys a challenge, as long as he can be successful. But in a situation where he may make mistakes in class and lose his status as a top performer in the eyes of others, his motivation takes a nosedive:

> If you're doing it [making mistakes] in front of people? I wouldn't really want to do it because when you make a mistake you kind of tend to get embarrassed and people will say, "No, you're doing it wrong. What are you doing?" and that tends to be

> embarrassing … Like if you're doing a group and there's tons of kids around you, when you ask for help then they know you're not understanding it. And then some people say that's really easy and they're like, "How are you not understanding it?" It makes you feel stupid, and then you get embarrassed.

Thus, as we saw in both the study of students making the transition to junior high school and in the praise studies, a fixed mindset framework may not hurt performance in conditions of relatively low challenge, where students' skills exceed the demands of the task and success is readily obtainable. But when the possibility of failure looms in a situation where ability may be evaluated, mindset makes all the difference in whether the student will be resilient and bounce back from difficulty or become helpless and founder (Blackwell et al. 2007; Mueller and Dweck 1998).

Changing Mindsets

Given that mindsets are influenced by messages from others, we wondered whether it would be possible to teach a growth mindset and improve students' motivation and performance as a result. To do this, we developed an eight-session workshop to teach students a growth mindset by teaching them about the brain and how it develops and grows stronger through learning. The workshop included an article, "You Can Grow Your Intelligence," images and explanations of how the brain works, and discussions about learning and growth, along with lessons on study skills. We randomly assigned seventh grade students in an urban middle school to either this growth mindset workshop or an alternative version that taught students about the brain and study skills, but without the information and activities focused on the malleable brain and developing intelligence. Both workshops were taught in the students' advisory sections by separate teams of researchers.

After the workshops, we asked the students' teachers, who were blind to the workshop condition of the students, to identify and describe those who had improved in motivation over the course of the year. Fully three-fourths of the students their teachers identified were from the growth mindset workshop group—a significant difference. Here is a typical comment a teacher made about the observed changes:

> Your workshop has already had an effect. L., who never puts in any extra effort and often doesn't turn in homework on time, actually stayed up late working for hours to finish an assignment early so I could review it and give him a chance to revise it. He earned a B+ on the assignment (he had been getting C's and lower).

We examined the students' performance in math over the course of the study. Prior to the intervention, the grades of students in both groups had very similar trajectories: they were declining from those obtained in sixth grade (the previous year) in the same school as the challenge level in the curriculum increased. And indeed, in the term following the intervention, the grades of the students in the control group continued to decline, but the grades of students in the growth mindset workshop reversed course, erasing the downturn (see Fig. 18.4) (Blackwell et al. 2007). Thus, the students who received instruction in the malleable brain and developing ability not only became more motivated, their math performance rebounded even as the curriculum continued to become more difficult over the course of seventh grade.

We have since developed and tested a blended-learning curriculum, *Brainology*®, based on this workshop, Brainology, and found that it promoted a similar shift in mindset, motivation, and performance. Here is how middle school students who completed this curriculum described the impact of learning about the malleable brain on their view of intelligence, effort, and challenge:

> You're not born dumb or born smart … Once you know how your brain works, it's much easier to control it—once you develop more neurons and connections, it's much easier to approach something that's harder for you.
>
> Probability, I was just like, "I don't get this at all. But I was just like, okay, I'm going to do this … I want to do this since it's so hard. I'm going to be like, Brain, you cannot just run away from this. I'm going to do this!

Other studies, with participants from middle school to college, have shown similar impacts

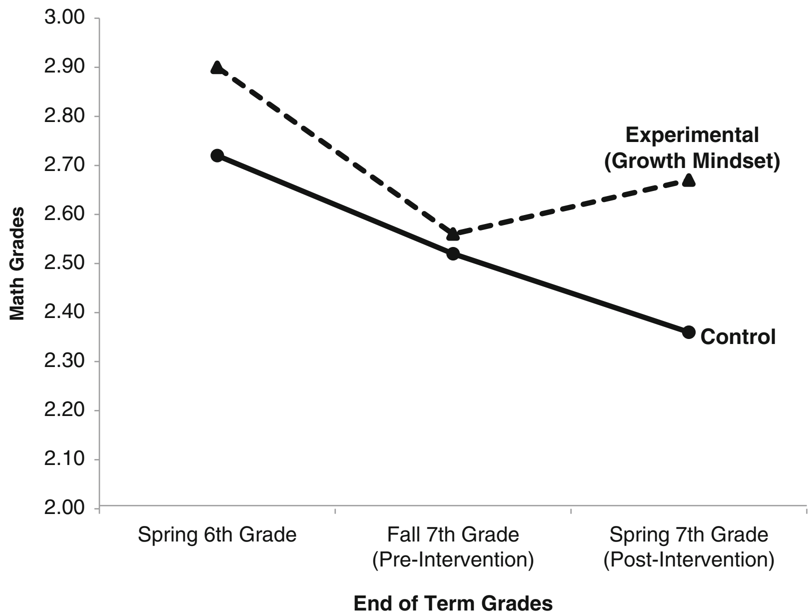


Fig. 18.4 Student achievement as a function of workshop group in Study 2 (Notes: All students' math grades had declined from the end of sixth grade to the end of the fall term in seventh grade, and there was no statistically significant difference between the two groups at either time point. Following the workshop (end of Spring seventh grade term), students in the intervention group (who learned about the malleable brain) rebounded, while the grades of students in the control group continued to decline ($p < .05$). Reprinted from Blackwell et al. 2007, p. 257. Reprinted with permission. Copyright 2007 from Society for Research in Child Development, Inc.)

of teaching a growth mindset on motivation and performance (e.g., Aronson et al. 2002; Good et al. 2003; Yeager et al. 2013). A series of new studies by David Yeager, David Paunesku, and their colleagues, for example, have found that even brief, one-time mindset interventions delivered online can lead to significant gains in achievement, including among students from traditionally underrepresented groups (Yeager et al. 2013). Thus, mindsets can be changed and motivation and achievement improved as a result.

Mindsets and the Achievement Gap

A particularly striking way that the social context can impact mindset and performance is embodied in the phenomenon of stereotype threat. Originally identified by Claude Steele and Joshua Aronson (Steele and Aronson 1995), stereotype threat occurs when students from a negatively stereotyped group (e.g., Black and Latino students in academics or female students in math and science) encounter a situation that puts them at risk of being judged in light of that stereotype and, potentially, of confirming it in the eyes of others. This concern with confirming a negative stereotype can interfere with thinking and motivation and, therefore, performance. For example, a female student taking a test of math ability given by a male administrator may worry that if she doesn't perform well, it will be seen as confirmation that females are not good at math, or a students of color taking the SAT may anticipate that their performance will be seen as reflecting on the intelligence of their race if they perform poorly.

The anticipation of such group-based negative evaluation can lead to a host of detrimental consequences, including negative thoughts (Cadinu et al. 2005; Keller and Dauenheimer 2003), anxiety (Marx and Stapel 2006), and physiological arousal that can reduce cognitive functioning (Blascovich et al. 2001; Krendl et al. 2008;

Osborne 2006, 2007; Vick et al. 2008), including working memory (Beilock et al. 2007; Schmader and Johns 2003) and attentional and behavioral control (Inzlicht et al. 2006; Smith and White 2002), especially on a challenging task (O'Brien and Crandall 2003; Stone and McWhinnie 2008), all of which can result in reduced performance and widening achievement gaps between groups (Beilock et al. 2007; Brown and Day 2006; Cadinu et al. 2005; Keller and Dauenheimer 2003; O'Brien and Crandall 2003; Schmader and Johns 2003; Steele and Aronson 1995). Thus, the preoccupation with ability and performance induced by stereotype threat can actually reduce both (at least temporarily) for students in contexts where the stereotype is relevant. However, when the task is defined as non-diagnostic of ability, the performance gap is narrowed (Aronson et al. 1999; Steele and Aronson 1995).

The negative effects of stereotype threat emerge for children as young as seven. In one set of studies (Hartley and Sutton 2013), boys underperformed on math, reading, and writing assessments when they held the belief that boys were "inferior at academics" and also when they were explicitly provided this message prior to working on the assessments. Interestingly, these effects disappeared when boys were instead told that there were no real differences in abilities between boys and girls.

Aronson and colleagues (2002) noted that the impact of stereotype threat on students—preoccupation with evaluation, anxiety during assessments, lower performance, and disidentification with academics—looked similar to the pattern observed in people with a fixed mindset concept of intelligence under conditions of challenge. Perhaps, they reasoned, teaching students to think of their abilities as malleable could buffer them against the negative effects of stereotype threat: knowing that they could always develop their ability, they would be less worried about whether they scored well on a particular test (Aronson et al. 2002). To test this, they taught students that their intelligence was malleable and then looked at their performance under conditions of stereotype threat. Teaching malleable intelligence was successful in increasing enjoyment and valuing of academics and academic performance, including GPA, among Black college students (Aronson et al. 2002). Similarly, in a study with middle school minority students, teaching about malleable intelligence raised achievement test scores in both reading and math and narrowed the achievement gap between male and female students (Good et al. 2003). Finally, explaining gender differences in math performance as the result of genetic factors rather than experience (Dar-Nimrod and Heine 2006) or effort (Thoman et al. 2008) reduced the performance of females on a math test consistent with the notion that malleable, experiential-based explanations of ability buffer students from the negative effects of stereotype threat.

These sets of studies highlight the powerful role that the psychological experience and context plays in shaping students' motivation, learning, and performance. In sum, the evidence strongly suggests that cognitive performance is the product of a synergistic relationship between individual aptitudes, beliefs, and preferences and influences from the environment, and that mindsets about intelligence are a critical part of this relationship.

Implications for Future Research

Many questions still remain to be answered regarding the interplay of mindsets and intellectual achievement. From a developmental perspective, we know too little about how early experience forges mindsets and the impact that it has on the development of talent and skills over a child's early years. In particular, research showing the emergence of the characteristic mindset patterns in young children, even before they develop a differentiated concept of intelligence, suggests that parental influence needs more investigation (Giles and Heyman 2003; Heyman et al. 2003; Smiley and Dweck 1994). The more general nature of these early mindsets raises a question of whether a more global growth or fixed mindset lies behind the much-investigated concept of ability: perhaps some children come to believe early on that people are generally fixed in their attributes, and this paradigm is later populated by more specific concepts such as intelligence, character, and the like.

How susceptible are mindsets to enduring change, and what is the minimum required intervention to drive meaningful change? Can lasting change be achieved through single point, targeted interventions teaching a growth mindset to students, or does it require a combination of malleable intelligence instruction with ongoing reinforcement through implicit messages, such as process praise? How much should we make these frameworks explicit and examined in order to transform them? Many studies indicate short-term effects of even brief growth mindset interventions (see, e.g., Yeager et al. 2013), but we do not yet know what it takes to achieve a permanent shift.

How complex and context dependent are mindsets? Note that, in much of the research literature, mindset has been treated as a global and categorical variable, contrasting fixed versus growth mindsets about general intelligence. However, it is possible to have different mindsets about different kinds of ability. For example, as shown in the studies using mindset interventions to reduce stereotype threat, some groups are vulnerable in specific subject areas, despite superior academic performance overall (e.g., females in math and science; Good et al. 2012). In addition, as we have seen, messages received from the learning environment, including our interactions with others, can impact mindset-related behavior significantly (Blackwell et al. 2007; Mueller and Dweck 1998) and can shape the neural underpinnings of the learning experience (e.g., Rodriguez et al. 2014). These findings suggest that many people may harbor mixed or flexible mindsets and rely on environmental cues to activate the one deemed most appropriate to the situation. Further research to learn more about the contextual factors that can activate different mindsets would be helpful in designing interventions and educational programs to support struggling learners.

Finally, what are the long-term consequences of holding a growth mindset for the development of one's abilities and talents? Walter Michel's impactful work has demonstrated how the tendency to delay immediate gratification during the preschool years is correlated with positive events across the lifespan (cf. Mischel et al. 2011). How does holding a growth mindset impact people over a lifetime? Research in the workplace, for example, finds that growth mindset in leadership roles ("leaders are made") can contribute to greater confidence and positive affect (Hoyt et al. 2012). Future research should continue to explore this possibility.

Implications for Educational Practice

Closing the Achievement Gap

While the debate over the nature of intelligence continues, our educational system, from K-12 to college, is grappling with the application of these concepts and measures to policy and practice, with often unintended consequences. In many schools, students are still ranked and tracked by ability based on prior achievement or their scores on assessments, exposing them to differing curricula and standards. Standing on achievement tests can be misapplied in practice as ability labels that lead both students and teachers to adopt a fixed mindset and lower expectations, which can then become a self-fulfilling prophecy. This emphasis on normative assessments and grading practices can make achievement appear to be a zero-sum game, with enormous implications due to competitive access to schools and higher education opportunities. The persistent achievement gap due to unequal educational opportunities and the psychological burden imposed by societal stereotypes for African American and Latino students and for females in math, science, and engineering still signal that a large number of our young people are laboring under identity-based fixed mindset conceptions of their ability that limit them in fulfilling their potential.

Indeed, even among individuals with both high math and verbal ability, women are less likely to pursue careers in science, technology, engineering, or math (e.g., Wang et al. 2013). These findings are especially striking because women in the Wang et al. (2003) study represented a greater percentage of those individuals with high scores on both math and verbal assessments. Recruiting students into these fields may pose challenges for additional reasons, one of these arising from the way in which these abilities are portrayed in popular media. For example, when students were presented with a (fictitious) newspaper articles conveying the biological "nature" of gender differences, readers were more likely to agree with gender stereotypes, whereas the opposite was true when social explanations were used to describe differences (Brescoll and LaFrance 2004). Explanations for group differences that rest upon biology implicitly convey a fixed mindset conception of ability and reinforce the stereotypes that can undermine motivation and achievement. Thus, both academic practices and their reflection in the popular culture can inadvertently constrain the performance of vulnerable students.

Educational Systems and Structures

The research shows that messages that highlight a person's ability, rather than their effort and process, reinforce a fixed mindset and often precipitate a helpless pattern when the person encounters difficulty. However, while individual teachers can change the way they praise and criticize students in their classroom, they also operate within a larger context of assessment and incentives that are not informed by this research. For example, most schools and districts still adhere to one-size-fits all curricula and standards with age and grade-level expectations, grade students in comparison to their peers on a rigid timeline, and rely on a small number of high-stakes tests to measure student and school success and to select students for access to future learning opportunities. For students who initially lack foundational skills and learning strategies, these policies may virtually ensure failure and undermine the focus on process and growth that are critical to promoting positive motivation.

Recent research suggests that motivation-related behaviors, such as "grit" and perseverance in pursuing goals, may be more predictive of success than IQ; for example, Angela Duckworth and colleagues have found that students who exhibited greater self-control earned higher grades, whereas students' IQ scores were not related to achievement (Duckworth and Seligman 2005). We know from a large body of research that teaching content in the absence of positive academic mindsets and basic learning skills often falls short and that programs aimed at changing students' learning behavior directly are less effective than interventions that change their mindsets. (See Farrington et al. 2012 for a review.) Yet the vast majority of our educational efforts are devoted to core subject curriculum and assessment, sometimes at the expense of teaching academic mindsets and foundational learning skills.

The evidence shows that intellectual development and performance are highly dependent on the interaction of environmental supports and individual effort and active engagement—and that practices that instill and nurture a growth mindset also promote and sustain effort, engagement, and achievement. What could we achieve if, rather than measuring and comparing students with one another, we focused on providing them with a solid foundation of self-efficacy and skills and then on creating opportunities for them to grow? With the knowledge we have gained from decades of research in the role of mindsets in achievement, we have the opportunity to provide the current generation of learners—and their teachers—with a solid foundation for future growth.

References

Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29–46.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping mindsets of intelligence. *Journal of Experimental Social Psychology, 38*, 113–125.

Bartels, M., Rietveld, M., Van Baal, G., & Boomsma, D. (2002). Genetic and environmental influences on the development of intelligence. *Behavior Genetics, 32*, 237–249.

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviations, and spillover. *Journal of Experimental Psychology: General, 136*, 256–276.

Binet, A. (1975). *Modern ideas about children.* (trans: Heisler, S.). Menlo Park: Suzanne Heisler. (originally work published in 1909).

Blackwell, L., Trzesniewski, K., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246–263.

Blascovich, J., Spencer, S. J., Quinn, D. M., & Steele, C. M. (2001). African-Americans and high blood pressure: The Role of stereotype threat. *Psychological Science, 12*, 225–229.

Borland, J. H. (2003). The death of giftedness. In J. H. Borland (Ed.), *Rethinking gifted education* (pp. 105–124). New York: Teachers College Press.

Borland, J. H. (2005). Gifted education without gifted children: The case for no conception of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed.). New York: Cambridge University Press.

Borland, J. H., & Wright, L. (2001). Identifying and educating poor and underrepresented gifted students. In K. A. Heller, F. J. Monks, R. J. Sternberg, & R. F. Subotnik (Eds.), *International handbook of research and development of giftedness and talent* (pp. 587–594). London: Pergamon Press.

Brehmer, Y., Westerberg, H., & Backman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience, 6*, 1–7.

Brescoll, V. L., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on gender differences. *Psychological Science, 15*(8), 515–521.

Breslau, N., Chilcoat, H., Susser, E., Matte, T., Liang, K., & Peterson, E. (2001). Stability and change in children's intelligence quotient scores: A comparison of two socioeconomically disparate communities. *American Journal of Epidemiology, 154*, 711–717. doi:10.1093/aje/154.8.711.

Brown, R. P., & Day, E. A. (2006). The difference isn't black and white: Stereotype threat and the race gap on Raven's advanced progressive matrices. *Journal of Applied Psychology, 91*, 979–985.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? *Psychological Science, 16*, 572–578.

Canivez, G., & Watkins, M. (1998). Long-term stability of the Wechsler intelligence scale for children—third edition. *Psychological Assessment, 10*, 285–291.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*(5), 703–722.

Cimpian, A., Mu, Y., & Erickson, L. C. (2012). Who is good at this game? Linking an activity to a social category undermines children's achievement. *Psychological Science, 23*(5), 533–541.

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64*, 5–31.

Darling-Hammond, L. (1995). Cracks in the bell curve: How education matters. *The Journal of Negro Education, 64*, 340–353.

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science, 314*, 435.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: Changes in grey matter induced by training. *Nature, 427*, 311–312.

Duckworth, A., & Seligman, M. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*, 939–944.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.

Erikson, K., Drevets, W., & Schulkin, J. (2003). Glucocorticoid regulation of diverse cognitive functions in normal and pathological emotional states. *Neuroscience and Biobehavioral Reviews, 27*, 233–246.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago: University of Chicago Consortium on Chicago School Research.

Flynn, J. (2007). *What is intelligence?* New York: Cambridge University Press.

Giles, J. W., & Heyman, G. D. (2003). Preschoolers' beliefs about the stability of antisocial behavior: Implications for navigating social challenges. *Social Development, 12*, 182–197.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology, 24*, 645–662.

Good, C., Rattan, A., & Dweck, C. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology, 102*(4), 700–717. doi:10.1037/a0026659.

Gunderson, E., Gripshoever, S., Romero, C., Dweck, C., Goldin-Meadow, S., & Levine, S. (2013). Parent praise to 1- to 3- year-olds predicts children's motivational frameworks 5 years later. *Child Development, 84*(5), 1526–1541. doi:10.1111/cdev.12064.

Hackman, D., & Farah, M. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Science, 13*(2), 65–73.

Hartley, B., & Sutton, R. (2013). A stereotype threat account of boys' academic underachievement. *Child Development, 84*, 1716–1733. doi:10.1111/cdev.12079.

Henderson, V. L., & Dweck, C. S. (1990). Motivation and achievement. In S. S. Feldman & G. R. Elliot (Eds.), *At the threshold: The developing adolescent* (pp. 308–329). Cambridge: Harvard University Press.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Hertzog, C., & Schaie, K. (1986). Stability and change in adult intelligence: 1. Analysis of longitudinal covariance structures. *Psychology and Aging, 1*, 159–171.

Heyman, G., Gee, C., & Giles, J. (2003). Preschool children's reasoning about ability. *Child Development, 74*, 516–534.

Hong, Y., Chiu, C., Dweck, C., Lin, D., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*, 588–599.

Hoyt, C., Burnette, J., & Inella, A. (2012). I can do that: The impact of implicit theories on leadership role model effectiveness. *Personality and Social Psychology Bulletin, 38*, 257–268. doi:10.1177/0146167211427922.

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science, 17*, 262–269.

Kaufman, S. (2013). *Ungifted: Intelligence redefined*. New York: Basic Books.

Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin, 29*, 371–381.

Kray, L. J., & Haselhuhn, M. P. (2007). Implicit negotiation beliefs and performance: Experimental and longitudinal evidence. *Journal of Personality and Social Psychology, 93*, 49–64.

Krendl, A. C., Richeson, J. A., Kelley, W. M., & Heatherton, T. F. (2008). The negative consequences of threat: A functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science, 19*, 168–175.

Lupien, S., McEwen, B., Gunnar, M., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews Neuroscience, 10*, 434–445.

Mackey, A. P., Hill, S., Stone, S., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science, 14*(3), 582–590. doi:10.1111/j.1467-7687.2010.01005.x.

Mackey, A., Whitaker, K., & Bunge, S. (2012). Experience-dependent plasticity in white matter microstructure: reasoning training alters structural connectivity. *Frontiers in Neuroanatomy, 6*, 1–9. doi:10.3389/fnana.2012.00032.

Mackey, A., Miller-Singley, A., & Bunge, S. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *Journal of Neuroscience, 33*(11), 4796–4803.

Mangels, J. A., Butterfield, B., Lamb, J., Good, C. D., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social-cognitive-neuroscience model. *Social Cognitive and Affective Neuroscience, 1*, 75–86.

Marx, D. M., & Stapel, D. A. (2006). Distinguishing stereotype threat from priming effects: On the role of the social self and threat-based concerns. *Journal of Personality and Social Psychology, 91*, 243–254.

Mischel, W., Ayduk, O., Berman, M., Casey, B., Gotlib, J., Kross, E., Teslovich, T., Wilson, N., Zayas, V., & Shoda, Y. (2011). Willpower over the lifespan: Decomposing self-regulation. *Social Cognitive and Affective Neuroscience, 6*(2), 252–256.

Moser, J., Schroder, H., Heeter, C., Moran, T., & Lee, Y. (2011). Mind your errors: Evidence for a neural mechanism linking growth mindset to adaptive post-error adjustments. *Psychological Science, 22*(12), 1484–1489.

Mueller, C. M., & Dweck, C. S. (1998). Intelligence praise can undermine motivation and performance. *Journal of Personality and Social Psychology, 75*, 33–52.

Naglieri, J., & Goldstein, S. (2009). *Practicioner's guide to assessing intelligence and achievement*. Hoboken: Wiley.

Neville, H., Stevens, C., Pakulak, E., Bell, T., Fanning, J., Klein, S., & Isbell, E. (2013). Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. *Proceedings of the National Academy of Sciences, 110*, 12138–12143. doi:10.1073/pnas.1304437110.

Nisbett, R. (2009). *Intelligence and how to get it: Why schools and cultures count*. New York: W.W. Norton & Company.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29*, 782–789.

Osborne, J. W. (2006). Gender, stereotype threat and anxiety: Psychophysiological and cognitive evidence. *Journal of Research in Educational Psychology, 8*, 109–138.

Osborne, J. W. (2007). Linking stereotype threat and anxiety. *Educational Psychology, 27*, 135–154.

Pomerantz, E. M., & Kempner, S. G. (2013). Mothers' daily person and process praise: Implications for children's theory of intelligence and motivation. *Developmental Psychology.*

Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok – not everyone can be good at math:" Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*, 731–737.

Rheinberg, F., Vollmeyer, R., & Rollett, W. (2000). Motivation and action in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation: Theory, research and application* (pp. 503–529). San Diego: Academic.

Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity, 1*, 313–336.

Rodriguez, S., Mangels, J., Guerra-Carrillo, B., & Higgins, T. (2014). *Frame of mind: Focusing students on performance or mastery yields a double dissociation of the neural processes predicting subsequent memory*. Manuscript in preparation.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*, 440–452.

Smiley, P. A., & Dweck, C. S. (1994). Individual differences in achievement goals among young children. *Child Development, 65*, 1723–1743.

Smith, J. L., & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles, 47*, 179–191.

Steele, C., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology, 44*, 445–452.

Thoman, D. B., White, P. H., Yamawaki, N., & Koishi, H. (2008). Variations of gender-math stereotype content affect women's vulnerability to stereotype threat. *Sex Roles, 58*, 702–712.

Vick, S. B., Seery, M. D., Blascovich, J., & Weisbuch, M. (2008). The effect of gender stereotype activation on challenge and threat motivational states. *Journal of Experimental Social Psychology, 44*, 624–630.

Wang, M., Eccles, J., & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science, 24*(5), 770–775.

Wolf, T. (1973). *Alfred Binet*. Chicago: University of Chicago Press.

Yeager, D., Paunesku, D., Walton, G., & Dweck, C. (2013). *How can we instill productive mindsets at scale? A review of the evidence and an initial R&D agenda*. [white paper]. Retrieved from: http://homepage.psy.utexas.edu/homepage/group/YeagerLAB/ADRG/Pdfs/Yeager%20et%20al%20R&D%20agenda%20-%206-10-13.pdf

Zatorre, R., Fields, R., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience, 15*(4), 528–536. doi:10.1038/nn.3045.

19 Creativity and Intelligence

Jonathan A. Plucker, Amber Esping, James C. Kaufman, and Maria J. Avitia

Creativity and Intelligence

Given the importance of the construct of human intelligence, it is not surprising that the relationship between intelligence and other constructs has been frequently studied. For example, psychologists maintain active research programs on the relationship between intelligence and working memory (Burgess et al. 2011), intelligence and bias in logical thinking (Stanovich, West, and Toplak 2013), and – increasingly – intelligence and health (Deary et al. 2010).

The potential association of creativity and intelligence has consistently received attention, with the early scientific studies of intelligence (e.g., Galton's *Hereditary Genius,* 1869) overlapping considerably with the constructs of creativity and talent development. Several seminal studies of creativity focus at least in part on intelligence (e.g., Barron 1963; Getzels and Jackson 1962; Wallach and Kogan 1965); indeed, one of the first leading creativity researchers, J. P. Guilford, began by studying intelligence (Guilford 1967). This attention is due, in part, to the fact that both constructs are integral to understanding talent and giftedness, and scholarship on gifted education has traditionally included studies in this area.

As Sternberg and O'Hara (1999) observed, the degree to which creativity and intelligence are related – and the nature of any such relationship – is "theoretically important, and its answer probably affects the lives of countless children and adults" (p. 269), and Plucker and Renzulli (1999) concluded that it is now not a matter of discovering *whether* intelligence and creativity are related, but rather of *how* they are related. Although those sentiments are nearly a generation old at this point, the amount of theoretical and empirical work on the topic has continued to grow (e.g., J. C. Kaufman and Plucker 2011; Kell et al. 2013; Kim et al. 2010), providing evidence that the topic will remain popular into the future.

Another reason for heightened attention is the growing popularity of the "21st century skills" in education systems around the globe. Many models and frameworks of these skills include constructs that are similar, if not identical, to intelligence and creativity (see Partnership for 21st Century Skills 2013). These models implicitly endorse the view that problem solving is a

J.A. Plucker (✉) • J.C. Kaufman
Department of Educational Psychology, Neag School of Education, University of Connecticut, 2131 Hillside Road, Unit 3007, Storrs, CT 06269-3007, USA
e-mail: jonathan.plucker@uconn.edu; James.kaufman@uconn.edu

A. Esping
College of Education, Texas Christian University, TCU Box 297900, Fort Worth, TX 76021, USA

M.J. Avitia
University of Connecticut Storrs, 456 Tolland Tpke Apt 25a, Willington, CT 06279, USA
e-mail: Maria.avitia@uconn.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_19,

key competency for both life and workplace success and that intelligence and creativity are integral parts of the ability to solve problems effectively and efficiently.

The specific relationship of intelligence and creativity to each other and to general problem solving or other positive outcomes has tremendous practical implications. If intelligence and creativity are very closely entwined, for example, then an overdependence on ability and achievement measures (such as the college entrance examinations) would be less worrisome. If the two constructs are only slightly related, then workplaces that value both intelligence and creativity would need to make sure to specifically assess both abilities. At a broader level, if intelligence and creativity are essential and distinct components for successful problem solving, then we need to make sure that our schools and workplaces are doing their best to nurture and support both abilities.

In What Ways Can Intelligence and Creativity Be Related?

Despite the importance of the creativity-intelligence relationship, a novice reading the literature for the first time could be forgiven for walking away bewildered: The breadth of theoretical treatments is vast, and there is surprisingly little empirical convergence. For example, threshold theory suggests intelligence is a necessary but not a sufficient condition of creativity (Barron 1969; Yamamoto 1964), certification theory focuses on environmental factors that allow people to display creativity and intelligence (Hayes 1989), and an interference hypothesis suggests that very high levels of intelligence may interfere with creativity (Simonton 1994; Sternberg 1996).

To complicate things further, intelligence and creativity are each time-consuming to assess well. Most studies use divergent thinking tests to measure creativity and group-based tests of *g* to measure intelligence (see Barron and Harrington 1981; Kim 2005). Although group IQ tests serve a strong purpose in research studies, they are not considered ideal for psychoeducational assessment (A. S. Kaufman and Lichtenberger 2006). In addition, most current IQ tests use the Cattell-Horn-Carroll (CHC; see Flanagan et al. 2007) model or other cognitive theories (e.g., Das et al. 1994) to give separate index scores in addition to producing an overall *g* score (A. S. Kaufman 2009). Plucker and Esping (2014) note that this bemusement is completely understandable, and they recommend focusing on the definitions of the constructs being studied in order to wade through and understand the wide range of conceptualizations. Although they were specifically talking about definitions of intelligence, the recommendation holds for definitions of creativity, and certainly (and perhaps most appropriately!) for comparisons of the two constructs.

Fortunately, Sternberg and O'Hara (1999) provide a framework for examining the construct definitions. Although Sternberg's approach is nearly 15 years old, a similarly productive alternative has not been subsequently proposed, and most, if not all, subsequent research fits into his framework. Due to the durability and usefulness of the model, we use it as the framework for this chapter. Sternberg's model suggests five possible relationships: creativity as a subset of intelligence, intelligence as a subset of creativity, creativity and intelligence as overlapping sets, creativity and intelligence as coincident sets, and creativity and intelligence as disjoint sets. In the following sections, we provide examples of each type of relationship. The last two categories, coincident and disjoint sets, are quite rare and are not described here.[1]

Creativity as a Subset of Intelligence

A number of psychometric theories include creativity, either explicitly or implicitly, as a part of intelligence, such as Guilford's Structure of the Intellect (SOI; 1967) model. He specifically included divergent thinking as a cognitive operation within the SOI model – many of the divergent thinking assessments used over the past 50

[1] That said, Nusbaum and Silvia (2011) offer some evidence of a coincident set perspective, which bears watching for future developments.

years originated from this model, and Renzulli (1973) developed a creativity curriculum based on the divergent thinking operation.

Gardner (1993), coming at the constructs from his developmental and qualitative perspective, has used Multiple Intelligence Theory to study creativity, implicitly suggesting that creativity is a subset of MI Theory. In Gardner's seminal book, *Creating Minds*, he used case studies of eminent creators to argue that people can excel creatively as a function of embodying different intelligences. For example, he highlighted Picasso (spatial intelligence), Freud (intrapersonal), Stravinsky (musical), Einstein (logical-mathematical), T. S. Eliot (linguistic), Martha Graham (bodily-kinesthetic), and Gandhi (interpersonal).

Another theory that includes creativity as a core component is Sternberg's (1996, 1997, 1999; Sternberg et al. 2008) theory of successful intelligence. This theory comprises three "subtheories": a *componential subtheory*, which relates intelligence to the internal world of the individual; an *experiential subtheory*, which relates intelligence to both the external and the internal worlds of the individual; and a *contextual subtheory*, which relates intelligence to the external world of the individual. The componential subtheory specifies the mental mechanisms responsible for planning, carrying out, and evaluating intelligent behavior. The experiential subtheory expands on this definition by focusing on those important behaviors that involve either adjustment to relative novelty, automatization of information processing, or both. The contextual subtheory defines intelligent behavior as involving purposeful adaptation to, selection of, and shaping of real-world environments relevant to one's life (Sternberg et al. 2008). The experiential subtheory is directly related to creativity. Sternberg's application of creativity assessments to admissions data increased prediction of college success beyond that obtained with standard admissions tests; in addition, ethnic-group differences were significantly reduced (Sternberg 2006, 2010, 2012; Sternberg and The Rainbow Project Collaborators 2005, 2006).

The Cattell-Horn-Carroll (CHC) model of intelligence also includes creativity as a subset of intelligence. The CHC model is a combination of the Cattell-Horn theory of fluid and crystallized intelligence (Horn and Cattell 1966, 1967) and Carroll's (1993) Three-Stratum Theory. Both the Cattell-Horn and Carroll models essentially take Spearman's (1904) *g*, with Horn and Cattell proposing two distinct forms of *g:* fluid intelligence (*Gf*), the ability to apply a variety of mental operations to solve novel problems, and crystallized intelligence (*Gc*), the breadth and depth of a person's accumulated knowledge of a culture and the ability to use that knowledge to solve problems (Horn and Cattell 1966; see also Horn and Blankson 2005).

Creativity was originally hypothesized to be strongly associated to *Gf* in the early stages of the Cattell-Horn *Gf-Gc* theory (Cattell and Butcher 1968). However, this connection is no longer explicitly part of the CHC theory, in part because empirical support has been mixed, with some studies supporting a link between Gc and creativity (Batey et al. 2009; Cho et al. 2010; Furnham and Chamorro-Premuzic 2006) and others providing evidence of a Gf-creativity link (Batey et al. 2010a, b; Nusbaum and Silvia 2011).

In current versions of CHC theory, creativity is primarily placed under the category of long-term storage and retrieval (*Glr*), the ability to store information in and fluently retrieve new or previously acquired information (e.g., concepts, ideas, items, names) from long-term memory. Indeed, *Glr* explicitly includes originality/creativity as one of its components (Flanagan et al. 2007). *Glr* has two parts, *learning efficiency* (learning and retaining new information) and *fluency* (generating many different possible solutions). Carroll (1993) proposed that these two parts were distinct abilities, labeled memory and learning and idea production, respectively. With the creation of CHC theory (McGrew 2005), they were combined into *Glr*. Carroll's idea production has long been associated with creativity (Schneider and McGrew 2012), most notably Guilford's (1967) operation of divergent production.

McGrew (2009) recently noted that "Some *Glr* narrow abilities have been prominent in creativity research (e.g., production, ideational fluency, or associative fluency)" (p. 6). In an otherwise detailed description of the model, this sentence is the only mention of creativity, originality, or divergent thinking. Fluid intelligence (*Gf*) is discussed in terms of its relationship to problem solving and coping with novel problems (both considered to be highly related to creativity). Nonetheless, in current discussions of the relationship of creativity to intelligence as presented by the CHC model, the emphasis is on *Glr*. Several of the narrow abilities incorporated into *Glr*, such as ideational fluency, associational fluency, and originality/creativity (obviously), have been specifically hypothesized to relate strongly to creativity (Kaufman et al. 2011).

There is a striking lack of studies on how creativity is empirically connected to *Glr*. Very recently, Silvia et al. (2013) conducted one of the few studies to examine this relationship. They examined how divergent thinking performance related to multiple verbal fluency tests (representing the lower-order *Glr* factors related to fluency). Silvia et al. found that the larger *Glr* factor had a significant effect on both fluency and originality in divergent thinking.

There is much more work to be done, and this gap is notable not only because of the theoretical link between *Glr* and creativity but also because of the existing work on the positive link between memory and creativity. Much of the work on creative cognition, from Wallas (1926) and later scholars (Finke et al. 1992; Mumford et al. 1991) to neuropsychological models (Bristol and Viskontas 2006; Gabora 2010), discusses cognitive processes that apply preexisting knowledge to new concepts. Mednick's (1962) theory of remote associations posited that the ability to organize and access ideas aided creativity by allowing for more remote associations between ideas.

Mednick's (1968) measure, the Remote Associates Test, is a common measure of creative problem solving that requires people to find a word that is associated with three other words (e.g., sleeping, bean, and trash are all connected to the word bag). Storm et al. (2011) found that when people take the Remote Associates Test, they are more likely to forget other common associated words. This *goal-directed forgetting* (Bjork et al. 1998) is linked to better creative problem solving.

An interesting extension of the CHC perspective was provided by Martindale (1999), who proposed that people who are creative are selective with their information processing speed. Early in the creative process, they focus on processing larger amounts of information, but as the problem in question becomes better understood, they shorten their attention span, thereby increasing their processing speed. Martindale's work is similar to Sternberg's (1981) hypothesis that brighter people spend more time in initial global planning so that they do not have to spend as much time in local planning later in the process.

A relatively new theory, the Dual Process Theory of Intelligence (S. B. Kaufman 2013), treats creativity as a subset of intelligence. S. B. Kaufman (2013), in an attempt to combine cognitive models with intelligence research, posits a two-factor model of intelligence, with one factor representing controlled cognition (type 1) – goal-directed thoughts and actions that include, but are entirely explained by, *g* – and one representing spontaneous cognition (type 2) – unconsciousness-related thoughts and actions that include implicit learning ability and daydreaming, among many other constructs.

In S. B. Kaufman's view, creativity results from the combination of type 1 and type 2 processes, with the role and importance of each type of process varying during any specific individual's creative process. S. B. Kaufman notes the pertinent research of Vartanian (2009) and colleagues (Vartanian et al. 2009), in which they provide evidence that creative individuals may have the ability to focus attention to varying degrees depending on the context in which their creative cognition is occurring and the type of creative activity undertaken.

Intelligence as a Subset of Creativity

In contrast, other researchers have hypothesized that intelligence is a part of creativity. Although this approach has received scant attention among intelligence researchers, systems models of creativity tend to emphasize the contribution of intelligence and related cognitive processes as one factor among many that influence the development of creativity. As systems theories have grown in prevalence in the social sciences, they have grown in importance within creativity (Kozbelt et al. 2010), making this conceptual category rather more popular than in the earlier years of the field.

One major creativity theory that includes intelligence is Sternberg and Lubart's (1996) "investment" theory, in which they use the metaphor of the stock market. The key to being creative, they argue, is to buy low and sell high with your ideas. In this model, a successful creator will have ideas that may be at first be unpopular or underappreciated yet will preserve and eventually persuade other people that his or her ideas are valuable. The creator will then know at what point to move on to pursue other ideas.

According to this theory, six main elements contribute to successful creativity: intelligence, knowledge, thinking styles, personality, motivation, and the environment. Intelligence contributes using three elements drawn from Sternberg's triarchic theory (1988, 1996; later expanded into the theory of successful intelligence as described earlier).

Another theory that casts creativity as being a blend of different abilities is Amabile's (1982, 1996) componential model of creativity. She argued that three variables were needed for creativity to occur: domain-relevant skills, creativity-relevant skills, and task motivation. Domain-relevant skills include knowledge, technical skills, and specialized talent (i.e., a creative writer should know basic grammar and styles). Creativity-relevant skills are more personal factors that have been associated with creativity. These can include tolerance for ambiguity, self-discipline, and risk-taking. Finally, Amabile highlights motivation (intrinsic or extrinsic) toward the task at hand. Intelligence would primarily occur at the domain-relevant skill level.

A third theory that accounts for multiple variables but takes a more domain-specific approach is the Amusement Park Theoretical Model (APT Model; Baer and Kaufman 2005; Kaufman and Baer 2004). In an amusement park, there are *initial requirements* (e.g., a ticket, a ride to the location) that apply to all areas of the park. Similarly, there are initial requirements that, to varying degrees, are necessary for creative performance in all domains. One such essential initial requirement is intelligence. There are then a series of subcomponents, *general thematic areas*, *domains*, and *microdomains*, that get more and more specific (e.g., social sciences to psychology to educational psychology). Different aspects of intelligence that are more or less important across these specific areas, for example, *Gc*, might be particularly important to a historian, whereas *Gf* might be essential for an engineer.

Overlapping Sets

Sternberg's third grouping conceptualizes intelligence and creativity as overlapping yet distinct constructs. Renzulli's (1978) three-ring conception of giftedness theorizes that giftedness – implicitly cast as high-level creative production – is caused by the overlap of high intellectual ability, creativity, and task commitment. From this perspective, creativity and intelligence are distinct constructs but overlap considerably under the right conditions. In a similar vein, the concept of planning abilities in the planning, attention-arousal, simultaneous and successive (PASS) theory appears to overlap with creativity (Naglieri and Kaufman 2001), and Plucker et al. (2004) view creativity and intelligence as related but distinct in their definition of creativity as "the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context" (p. 90).

Threshold Theory. Traditional research has argued for a *threshold theory*, in which creativity and intelligence are positively, if moderately, correlated up until an IQ of approximately 120; in people with higher IQs, the two constructs show little relationship (e.g., Fuchs-Beauchamp et al. 1993; Getzels and Jackson 1962). Sternberg places this perspective in the overlapping sets category. This view is so common as to be considered part of the conventional wisdom about creativity, intelligence, and giftedness.

Yet some recent work has called into question the presence of a threshold (whether lower or higher than a 120 IQ). For example, Preckel et al. (2006) studied *Gf* and creativity (as measured through divergent thinking tests) and found modest correlations across all levels of intellectual abilities. In a meta-analysis of 21 studies, Kim (2005) found virtually no support for the threshold theory; there were small positive correlations between measures of ability and measures of creativity and divergent thinking.

Many of the early studies that formed the basis of threshold theory would now be considered quite dated. Any study conducted before the 1970s would have (obviously) used measures of intelligence that do not reflect current theory. Other studies have defined creativity rather narrowly as being only divergent thinking or the ability to generate multiple ideas in response to a single prompt.

Fortunately, researchers have begun to address these limitations, with interesting results. For example, Sligh et al. (2005) used a contemporary, individually administered IQ test (Kaufman Adolescent and Adult Intelligence Test; A. S. Kaufman and Kaufman 1993) and a creative invention task (in which people use shapes to create product objects and then name and describe their invention; see Finke 1990). By assessing both *Gc* and *Gf*, they were able to show moderate, positive correlations between *Gc* and creativity (i.e., similar to previous studies); however, intelligence and creativity were significantly correlated for the high-IQ group, which was not the case for people with average intelligence scores – an opposite pattern than predicted by threshold theory.

In a similar line of research (but with much different participants), the Study of Mathematically Precocious Youth has been following a cohort of students from late childhood/early adolescence into adulthood. These students all scored in the top 1 % on college entrance examinations before the age of 13, so they are a very bright group. Park et al. (2007, 2008) found that, within this very intelligent group, intellectual talent was highly correlated with educational attainment. This by itself is not surprising, but they also found that a range of indicators of adult creative accomplishments (e.g., patents, publications, awards) was also correlated with intelligence. Wai et al. (2005), looking at the same population, found that differences in SAT scores – even within such an elite group – predicted creative accomplishments 20 years later. These results, emerging as they do from studies that address the limitations of previous research, raise serious doubts about the threshold effect, and they firmly reflect a belief in the overlapping sets perspective.

Recent research by Beaty and Silvia (2012) provides potential insight into the mechanisms at play here. They had college students complete divergent thinking tasks over a 10-min period, and they unsurprisingly found that participants reported more creative ideas as they progressed throughout the session, as time marched on. However, they also administered a *Gf* measure, and – surprisingly – the higher the *Gf* score, the flatter the slope of the creativity-time curve. This means that the most intelligent people in their sample did not come up with appreciably more creative ideas over time, rather providing fairly creative ideas from the beginning to the end of the session. Less intelligent participants had increasingly steep slopes, meaning that they definitely were more creative as time progressed. This study raises the possibility that there are underlying cognitive mechanisms behind recent observations that intelligence and creativity are correlated even at high levels of intelligence and that those mechanisms may be a combination of executive processes related to information retrieval and manipulation and associative processes that involve activation of various parts of one's cognitive schema.

One recent study, by Jauk et al. (2013), uses sophisticated statistical techniques to support the threshold effect, if not the traditional cutoff level. They used segmented regression analysis and found that the threshold level varied based on the creativity score used. Ideational fluency's threshold was at an IQ of 86 (which is astoundingly lower than the more traditional 120); if the top two originality scores were used, the threshold was 104.

Conclusion

Each of the five possible relationships in Sternberg's framework enjoys at least some empirical support (Sternberg and O'Hara 1999; Kaufman and Plucker 2011), but the difficulty in interpreting empirical results illustrates the problems associated with reaching a consensus on the validity of any of these five relations. For example, Haensly and Reynolds (1989) believe that Mednick's (1962) Association Theory supports the creativity as a subset of intelligence position, yet Sternberg and O'Hara (1999) feel that this body of work supports the overlapping sets position. If Gardner's work with creativity had come before his work with MI Theory, we would probably be arguing that his efforts fall within the intelligence as a subset of creativity category.

As Plucker and Esping (2014) note, definitions are critically important when dealing with psychological constructs, as the way in which each construct is conceptualized and assessed will have a significant impact on any empirical results when comparing two or more constructs. As this chapter has showed, the range of creativity and intelligence definitions makes the complexity of possible intelligence-creativity relationships unsurprising. Few people believe creativity and intelligence are completely unrelated, but the nature of any relationship – even with so much research already conducted – is an open question.

References

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*, 997–1013.

Amabile, T. M. (1996). *Creativity in context: Update to "The Social Psychology of Creativity"*. Boulder: Westview Press.

Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The Amusement Park Theoretical (APT) Model of creativity. *Roeper Review, 27*, 158–163.

Barron, F. (1963). *Creativity and psychological health*. Princeton: D. Van Nostrand Company.

Barron, F. (1969). *Creative person and creative process*. New York: Holt, Rinehart & Winston.

Barron, F., & Harrington, D. M. (1981). Creativity, intelligence, and personality. *Annual Review of Psychology, 32*, 439–476.

Batey, M., Chamorro-Premuzic, T., & Furnham, A. (2009). Intelligence and personality as predictors of divergent thinking: The role of general, fluid, and crystallized intelligence. *Thinking Skills and Creativity, 4*, 60–69.

Batey, M., Chamorro-Premuzic, T., & Furnham, A. (2010a). Individual differences in ideational behavior: Can the big five and psychometric intelligence predict creativity scores? *Creativity Research Journal, 22*, 90–97.

Batey, M., Furnham, A., & Safiullina, X. (2010b). Intelligence, general knowledge and personality as predictors of creativity. *Learning and Individual Differences, 20*, 532–535.

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 309–319.

Bjork, E. L., Bjork, R. A., & Anderson, M. C. (1998). Varieties of goal-directed forgetting. In J. M. Golding & C. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103–137). Hillsdale: Erlbaum.

Bristol, A. S., & Viskontas, I. V. (2006). Dynamic processes within association memory stores: Piecing together the neural basis of creative cognition. In J. C. Kaufman & J. Baer (Eds.), *Creativity and reason in cognitive development* (pp. 60–80). New York: Cambridge University Press.

Burgess, G. C., Gray, J. R., Conway, A. R., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General, 140*, 674–692.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Cattell, R. B., & Butcher, H. (1968). *The prediction of achievement and creativity*. Indianapolis: Bobbs-Merrill.

Cho, S. H., Nijenhuis, J. T., van Vianen, N. E. M., Kim, H.-B., & Lee, K. H. (2010). The relationship between diverse components of intelligence and creativity. *Journal of Creative Behavior, 44*, 125–137.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Boston: Allyn & Bacon.

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53–79.

Finke, R. (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale: Erlbaum.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Flanagan, D. P., Ortiz, S., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd ed.). New York: Wiley.

Fuchs-Beauchamp, K. D., Karnes, M. B., & Johnson, L. J. (1993). Creativity and intelligence in preschoolers. *Gifted Child Quarterly, 37*, 113–117.

Furnham, A., & Chamorro-Premuzic, T. (2006). Personality, intelligence and general knowledge. *Learning and Individual Differences, 16*, 79–90.

Gabora, L. (2010). Revenge of the 'neurds': Characterizing creative thought in terms of the structure and dynamics of human memory. *Creativity Research Journal, 22*, 1–13.

Galton, F. (1869). *Hereditary genius*. London: Macmillan and Company.

Gardner, H. (1993). *Creating minds*. New York: Basic Books.

Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students*. New York: Wiley.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Haensly, P. A., & Reynolds, C. R. (1989). Creativity and intelligence. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 111–132). New York: Plenum.

Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 135–145). New York: Plenum.

Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 41–68). New York: Guilford.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57*, 253–270.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107–129.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. C. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence, 41*, 212–221.

Kaufman, A. S. (2009). *IQ Testing 101*. New York: Springer.

Kaufman, S. B. (2013). *Ungifted: Intelligence redefined*. New York: Basic Books.

Kaufman, J. C., & Baer, J. (2004). The Amusement Park Theoretical (APT) Model of creativity. *Korean Journal of Thinking and Problem Solving, 14*, 15–25.

Kaufman, A. S., & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines: American Guidance Service.

Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3rd ed.). New York: Wiley.

Kaufman, J. C., & Plucker, J. (2011). Creativity and intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 771–783). New York: Cambridge.

Kaufman, J. C., Kaufman, S. B., & Lichtenberger, E. O. (2011). Finding creativity on intelligence tests via divergent production. *Canadian Journal of School Psychology, 26*, 83–106.

Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychological Science, 24*, 1831–1836.

Kim, K. H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education, 16*, 57–66.

Kim, K. H., Cramond, B., & VanTassel-Baska, J. (2010). The relationship between creativity and intelligence. In J. C. Kaufman & R. J. Sternberg (Eds.), *Cambridge handbook of creativity* (pp. 395–412). New York: Cambridge University Press.

Kozbelt, A., Beghetto, R. A., & Runco, M. A. (2010). Theories of creativity. In R. J. Sternberg & J. C. Kaufman (Eds.), *The Cambridge handbook of creativity* (pp. 20–47). New York: Cambridge.

Martindale, C. (1999). Biological bases of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 137–152). New York: Cambridge University Press.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 136–181). New York: Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 69*, 220–232.

Mednick, S. A. (1968). The remote associates test. *Journal of Creative Behavior, 2*, 213–214.

Mumford, M. D., Mobley, M. I., Uhlman, C. E., Reiter-Palmon, R., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal, 4*, 91–122.

Naglieri, J. A., & Kaufman, J. C. (2001). Understanding intelligence, giftedness, and creativity using PASS theory. *Roeper Review, 23*, 151–156.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence, 39*, 36–45.

Park, G., Lubinski, D., & Benbow, C. P. (2007). Contrasting intellectual patterns predict creativity in the arts and sciences. *Psychological Science, 18*, 948–952.

Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability differences among people who have commensurate

degrees matter for scientific creativity. *Psychological Science, 19*, 957–961.
Partnership for 21st Century Skills. (2013). *Framework for 21st century learning*. Accessed at http://www.p21.org/our-work/p21-framework
Plucker, J., & Esping, A. (2014). *Intelligence 101*. New York: Springer.
Plucker, J., & Renzulli, J. S. (1999). Psychometric approaches to the study of human creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 35–60). New York: Cambridge University Press.
Plucker, J., Beghetto, R. A., & Dow, G. (2004). Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*, 83–96.
Preckel, F., Holling, H., & Wiese, M. (2006). Relationship of intelligence and creativity in gifted and non-gifted students: An investigation of threshold theory. *Personality and Individual Differences, 40*, 159–170.
Renzulli, J. S. (1973). *New directions in creativity*. New York: Harper & Row.
Renzulli, J. S. (1978). What makes giftedness? Reexamining a definition. *Phi Delta Kappan, 60*, 180–261.
Schneider, W. J., & McGrew, K. (2012). The Cattell-Horn-Carroll model of intelligence. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). New York: Guilford.
Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence, 41*, 328–340.
Simonton, D. K. (1994). *Greatness: Who makes history and why*. New York: Guilford Press.
Sligh, A. C., Conners, F. A., & Roskos-Ewoldsen, B. (2005). Relation of creativity to fluid and crystallized intelligence. *Journal of Creative Behavior, 39*, 123–136.
Spearman, C. (1904). "General intelligence", objectively determined and measured. *American Journal of Psychology, 15*, 201–293.
Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science, 22*, 259–264.
Sternberg, R. J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology, 73*, 1–16.
Sternberg, R. J. (1988). A three-facet model of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 125–147). New York: Cambridge University Press.
Sternberg, R. J. (1996). *Successful intelligence*. New York: Simon & Schuster.
Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.
Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology, 3*, 292–316.
Sternberg, R. J. (2006). Creating a vision of creativity: The first 25 years. *Psychology of Aesthetics, Creativity, and the Arts, S*, 2–12.
Sternberg, R. J. (2010). Teaching for creativity. In R. A. Beghetto & J. C. Kaufman (Eds.), *Nurturing creativity in the classroom* (pp. 394–414). New York: Cambridge University Press.
Sternberg, R. J. (2012). The assessment of creativity: An investment-based approach. *Creativity Research Journal, 24*, 3–12.
Sternberg, R. J., & The Rainbow Project Collaborators. (2005). Augmenting the SAT through assessments of analytical, practical, and creative skills. In W. Camara & E. Kimmel (Eds.), *Choosing students: Higher education admission tools for the 21st century* (pp. 159–176). Mahwah: Lawrence Erlbaum Associates.
Sternberg, R. J., & Lubart, T. I. (1996). *Defying the crowd*. New York: Free Press.
Sternberg, R. J., & O'Hara, L. A. (1999). Creativity and intelligence. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 251–272). Cambridge, MA: Cambridge University Press.
Sternberg, R. J., & The Rainbow Project Collaborators. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical and creative skills. *Intelligence, 34*, 321–350.
Sternberg, R. J., Kaufman, J. C., & Grigorenko, E. L. (2008). *Applied intelligence*. Cambridge: Cambridge University Press.
Storm, B. C., Angello, G., & Bjork, E. L. (2011). Thinking can cause forgetting: Memory dynamics in creative problem solving. *Journal of Experimental Psychology Learning Memory and Cognition, 37*, 1287–1293.
Vartanian, O. (2009). Variable attention facilitates creative problem solving. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 57–59.
Vartanian, O., Martindale, C., & Matthews, J. (2009). Divergent thinking ability is related to faster relatedness judgments. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 99–103.
Wai, J., Lubinski, D., & Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youths: An age 13 to age 33 longitudinal study. *Journal of Educational Psychology, 97*, 484–492.
Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity-intelligence distinction*. New York: Holt, Rinehart and Winston.
Wallas, G. (1926). *The art of thought*. New York: Harcourt Brace.
Yamamoto, K. (1964). Creativity and sociometric choice among adolescents. *Journal of Social Psychology, 64*, 249–261.

Part IV

Assessment of Intelligence

20 Hundred Years of Intelligence Testing: Moving from Traditional IQ to Second-Generation Intelligence Tests

Jack A. Naglieri

"Do not go where the path may lead, go instead where there is no path and leave a trail."

–Ralph Waldo Emerson

Context

April 6, 1917, is remembered as the day the United States entered World War I. On that same day a group of psychologists held a meeting in Harvard University's Emerson Hall to discuss the possible role they could play with the war effort (Yerkes 1921). The group agreed that psychological knowledge and methods could be of importance to the military and utilized to increase the efficiency of the Army and Navy personnel. The group included Robert Yerkes, who was also the president of the American Psychological Association. Yerkes made an appeal to members of APA who responded by providing a group of psychologists to assist with the war effort. Members from APA were joined by psychologists of the National Research Council, the National Academy of Sciences, and the American Association for the Advancement of Science, and a number of committees were organized to develop effective measures of ability.

One group of psychologists whose task was to begin identifying possible tests met at the Training School in Vineland, New Jersey, on May 28. The committee considered many types of group tests and several that Arthur S. Otis developed when working on his doctorate under Lewis Terman at Stanford University. The goal was to find tests that could efficiently evaluate a wide variety of men, be easy to administer in the group format, and be easy to score. By June 9, 1917, the materials were ready for an initial trial. Men who had some educational background and could speak English were administered the verbal and quantitative (Alpha) tests and those that could not read the newspaper or speak English were given the Beta tests (today described as nonverbal).

The Alpha tests were designed to measure general information (e.g., how many months are there in a year?), common sense (e.g., why do we use stoves?), and verbal knowledge (synonyms/antonyms, verbal analogies, number series, disarranged sentences) (e.g., determine if a group of words could be sequenced to make a true statement) and to determine how well the examinee could follow verbal directions (e.g., draw a line from circle 3 to circle 6). The Beta tests included completion of a maze, construction of a design using blocks, number symbol association, identifying what is missing in a picture, and copying geometric shapes. Why two tests? Because, as Yoakum and Yerkes (1920) clearly stated, the Alpha test was an appropriate measure of intelligence for men who could read and write English

J.A. Naglieri, Ph.D. (✉)
University of Virginia, 6622 White Post Road, Centreville, VA 20121, USA
e-mail: jnaglieri@gmail.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_20,

sufficiently. The Beta tests were intended for those who had difficulty reading or spoke English poorly as well as those who were illiterate or not able to understand English (p. 51). The testing procedures ensured that men "who fail in alpha are sent to beta in order that injustice by reason of relative unfamiliarity with English may be avoided (Yoakum & Yerkes, p. 19)."

By July 7, 1917, the initial development of tests to measure intellectual ability and a study involving 400 cases was completed. The data obtained from testing sites in Indiana, Tennessee, and Syracuse and Brooklyn, New York, was shipped to the statistical unit in New York City for data analysis. Statistical analysis was directed by E. L. Thorndike with assistants A. S. Otis and L. L. Thurstone. The report of their analysis completed on July 20, 1917, showed that the tests could be appropriate to (a) "aid in segregating and eliminating the mentally incompetent, (b) classify men according to their mental ability, and (c) assist in selecting competent men for responsible positions" (Yerkes 1921, p. 19). Thus, July 20, 1917, could be considered the birth date of the verbal, quantitative, and nonverbal IQ test format which will be called traditional IQ in the remainder of this chapter.

By early 1918 a group of about 260 men trained in the Medical Corps School for Military Psychology began using the Army Alpha and Beta tests. Among them was the 22-year-old David Wechsler (1896–1981) who arrived at Fort Logan, Texas, in August that year. Wechsler, like Yerkes, who wrote in the Introduction to the *Psychological Examining in the United States Army* (1921), noticed "the educational, industrial, and significance of the methods [Alpha and Beta] (p. 5)". Wechsler's adaptation some 20 years later resulted in the Wechsler-Bellevue Intelligence Scale (1939) which also contained verbal and quantitative (the Alpha) and performance (the Beta) test questions. These tests would have been familiar to the founders of the company that ultimately published Wechsler's test (Wechsler 2005), because many years earlier (in 1917), R. S. Woodworth was the chief examiner at the Brooklyn testing site collecting initial data on the Army Alpha and Beta, Thorndike was responsible for the analysis of those data, and R. Cattell was initially on the original National Research Council meeting that led to the development of the measures.

The transition from Army Alpha and Beta to Wechsler IQ as we know it today is the result of the work of those psychologists who first met on April 6, 1917, and those that developed, validated, and used the Army Alpha and Beta tests. It would likely have been well beyond the expectations of Major Lewis M. Terman, Captain Edwin G. Boring, and related personnel including R. S. Woodworth, E. L. Thorndike, A. S. Otis, and L. L Thurstone that the work they did would become the most widely used approach to measuring intelligence in history. And in addition, their work (described in the book *Army Mental Tests* by Yoakum and Yerkes 1920) would define the structure of individual- (e.g., Wechsler Scales) as well as group-administered (Otis-Lennon 1979) IQ tests for the next 100 years.

The evolution of traditional IQ tests from their birth in 1917 has been defined by the many revisions of Wechsler and Otis-Lennon tests, the most current version of the latter being number 8 and the forthcoming fifth edition of the Wechsler Intelligence Scale for Children (WISC-V, Wechsler 2014). Ironically, there have been many revisions of these tests long after the authors died (Otis in 1964 and Wechsler in 1981). What has evolved is a body of knowledge about these tests and how to interpret them. For example, the first book on the development and interpretation of the Wechsler Scales was the 1939 publication of *Wechsler's Measurement and Appraisal of Adult Intelligence*. The fifth edition of that book was published (authored by Matarazzo) in 1976. Shortly thereafter, Alan Kaufman's 1979 book, *Intelligence Testing with the WISC-R,* provided practitioners with a wealth of information about both psychometrically sound and clinically relevant interpretation methods. This was followed by *Intelligence Testing with the WISC-III* (Kaufman 1994) and *WISC-IV* (Flanagan and Kaufman 2004). Other authors have also provided books on how to extract information about Wechsler's scales (e.g., Prifitera and Saklofske 1998; Weiss et al. 2006). All this effort has been focused on ways to interpret a test based on measures first assembled in 1917.

Despite the widespread use and acceptance of the traditional verbal/quantitative (Alpha) and nonverbal (Beta) IQ approach to intelligence, it is important to ask if this approach meets the demands of psychology and education today. Are the tests appropriate for diverse populations? Does the information traditional IQ tests yield assist in planning instruction and academic interventions? How well do these tests evaluate the intellectual component of a specific learning disability? Perhaps most importantly, should we continue to use IQ tests which were devoid of a theoretical foundation from the time they were first introduced more than 100 years ago (Naglieri and Kaufman 2008)? The purpose of this chapter is to provide answers to these questions and evidence for an alternative to traditional IQ based on brain function. The short answer is yes, our understanding of intelligence has evolved and better ways to measure it are now available.

Starting Over

The evolution of intelligence tests was stimulated by the publication of two *second-generation* ability tests. First was the publication of the *Kaufman Assessment Battery for Children* (Kaufman and Kaufman 1983). This test was based on a blend of perspectives about what intelligence may be. What the Kaufman's did that was most important was to tie a test of intelligence with a two-dimensional model of ability conceptualized within a cognitive processing context. Another very important aspect of the KABC and the second edition (K-ABC-II, Kaufman and Kaufman 2004) was the shift away from organizing their test based on the verbal, quantitative, and nonverbal content of the test questions. Instead, tests with verbal content were placed on an achievement scale and ability was measured using the sequential and simultaneous processing scales. This shift in emphasis from test content to the process needed to solve the problem put more emphasis on the cognitive activities of the examinee and resulted in a dramatic change in conceptualization of ability.

The Kaufmans' emphasis on the need for a view of ability was also important. They recognized the fact that traditional IQ lacked a theory, just as Pintner (1923) noted when he wrote "we did not start with a clear definition of general intelligence… [but] borrowed from every-day life a vague term implying all-round ability and… we [are] still attempting to define it more sharply and endow it with a stricter scientific connotation" (p. 53)." The Kaufman's stressed the point that a test of intelligence should be built upon a theory of intelligence.

Another *second-generation* ability test which was published by Naglieri and Das in 1977 further stimulated the evolution of the field. This test, called the *Cognitive Assessment System*, and the more recent CAS2 (Naglieri et al. 2014), was developed on a specific theory derived from the integration of cognitive and neuropsychological research described by A.R. Luria (1963, 1966, 1973, 1980). The theory is called PASS which stands for planning, attention, simultaneous, and successive neurocognitive abilities. Planning is the ability to perform complex decision making (related to the frontal lobes); attention is the ability to focus thinking and resist distractions (related to the brain stem); simultaneous processing ability is needed for understanding interrelationships (occipital/parietal area); and successive processing ability is used whenever sequencing is required (temporal lobes). The PASS theory was initially presented in the book *Assessment of Cognitive Processes: The PASS Theory of Intelligence* (Das et al. 1994) and elaborated by Naglieri (1999). More recently, the validity of the PASS theory as measured by the CAS and CAS2 is summarized in several resources (e.g., Naglieri and Das 2001; Naglieri 2012; Naglieri and Goldstein 2011; Naglieri and Otero 2011; Naglieri and Conway 2009; Naglieri et al. 2012, 2014).

When the KABC and CAS were introduced, these *second-generation* tests marked a change in the way intelligence was conceptualized and, just as importantly, measured. These two tests are most alike in their emphasis on measuring ability separately from academic skills. That is, they move away from the verbal/quantitative (Alpha) and nonverbal (Beta) organization of questions. The authors of these tests recognized that even though verbal and nonverbal are often described as types of intelligence, the authors of the Alpha and Beta never intended to measure two abilities, and neither did Wechsler. Wechsler's view of intelligence was not that verbal and nonverbal were two types of intelligence. Despite the fact that his tests yielded verbal and performance (nonverbal) IQ scores, Wechsler (1958) wrote: "the subtests are different measures of intelligence, not measures of different kinds of intelligence" (1958, p. 64). Boake (2002) noted that Wechsler viewed verbal and nonverbal (also labeled as performance) tests as equally valid measures of intelligence. Similarly, Naglieri (2003) clarified that the terms verbal and nonverbal "refer to the content of the test, not a type of ability" (p. 2). Moreover, Wechsler argued that nonverbal tests help to "minimize the overdiagnosing of feeble-mindness that was, he believed, caused by intelligence tests that were too verbal in content… and he viewed verbal and performance tests as equally valid measures of intelligence and criticized the labeling of performance [nonverbal] tests as measures of special abilities" (Boake 2002, p. 396).

Elimination of Army Alpha-type questions from a measure of intelligence was a bold move by these authors of *second-generation* ability tests and one that raised two important questions: (1) "How similar are verbal and quantitative test questions on an IQ test to an achievement test?" and (2) "Can verbal and quantitative tasks be taken out of a measure of intelligence without losing validity?"

Do Verbal and Quantitative Test Questions Differ from Achievement Test Questions?

It would seem reasonable that an IQ test should measure something different than an academic achievement test, but this is not the case. The verbal and quantitative portions of traditional IQ tests are remarkably similar to questions found in the achievement tests used to measure knowledge and skills. For example, verbal questions are found on both traditional IQ tests and measures of achievement. All traditional IQ tests include a measure of word knowledge just as measures of achievement do. Children are required to define a word like "bat" on subtests included in the SB-5 and WISC-IV *intelligence* tests and the WJ-III *achievement* test. The WJ-III Cognitive battery contains a Verbal Comprehension subtest that has the item like "tell me another word for small" and the WJ-III Achievement battery contains a Reading Vocabulary question: "tell me another word for little." In addition, an item on the WJ-III Reading Vocabulary achievement test is "Tell me another word for (examiner points to the word big)," and in a Cognitive battery the examiner asks something like: "Tell me another word for tiny." Additionally, the WJ-III Cognitive battery Verbal Comprehension contains Picture Vocabulary items, and the WJ-III Achievement battery includes Picture Vocabulary items, some of which are the same. The WJ-III Cognitive tests also require the subject to name as many examples as possible from a given category in a 1-min time period and the same question appears on the WIAT-II Oral Expression achievement test. These examples do not comprise a complete list of item overlap but do represent the most strikingly similar questions.

Tests of intelligence and achievement also include arithmetic test questions. For example, the oldest intelligence test, now in its fifth edition is the *Stanford-Binet 5* (SB-5; Roid, 200×), contains Quantitative Reasoning items, one of which requires the child to calculate the total number of stars on a page (e.g., two stars in one box plus four in a second box plus one in a third box). Similarly, the *Wechsler Intelligence Scale for Children – Fourth Edition* (WISC-IV; Wechsler 2003) arithmetic subtest requires the child to count the number of butterflies pictured on a page. Although the scores these test items yield are used to determine the child's level of intelligence, very similar items appear on the *Wechsler Individual Achievement Test* (WIAT-II, Wechsler, 200×). On that test, for example, a Numerical Operations subtest item requires the child to determine the total number of marbles shown (e.g., 3 plus 5). Similarly, a *Woodcock-Johnson Tests of Achievement* (WJ-III; Woodcock et al. 2001) Applied Problems subtest item asks the child to count the number of crayons pictured on the stimulus book (e.g., 4). Moreover, a SB-5 Quantitative Reasoning item requires the child to complete a simple math problem (e.g., 3+2=5) just as the WJ-III Math Fluency (e.g., 5+2=?) and the WIAT-II Numerical Operations (e.g., 2+2=?) achievement tests do. There is lack of distinction between the arithmetic questions on these tests of achievement and intelligence, yet the interpretations of the scores each test yields are considerably different. In one instance the score is used to determine level of math achievement, but in the other the scores are used to determine level of intelligence. The same overlap in content is found for verbal tests.

The use of items with similar content across achievement and IQ tests is alarming for several reasons. First, because the correlation between IQ, especially verbal sections of IQ tests, and achievement test scores has been considered a source of evidence for the validity of IQ tests, the correlations between IQ tests with verbal/quantitative items and achievement tests should be considered overestimates of the relationship between ability and achievement. Moreover, the authors and/or publishers of IQ tests should justify how similar questions can be used across supposedly different constructs and how very different interpretations (e.g., achievement vs. intelligence) can be made given the similarity of item content. Second, the obvious achievement content must be justified by those that use these IQ tests when assessing culturally and/or linguistically diverse children and especially Hispanics who now constitute the largest minority group in the United States. This group is particularly at risk of being misdiagnosed because they often have parents with limited educational background and/or limited English language skills (Ramirez and de la Cruz 2002) which reduces the opportunity to acquire the knowledge of words (Hart and Risley 1995).

Can Verbal and Quantitative Tasks Be Taken Out of a Measure of Intelligence Without Losing Validity?

Explaining current academic successes and failures and predicting achievement over time is a critically important role an intelligence test can play. Having IQ test questions that measure very similar content to achievement tests enhances the predictive validity of these measures but at a cost to those with limited educational backgrounds (recall that the Beta tests were used to measure intelligence fairly for those with limited familiarity with English). The question remains, however, can *second-generation* intelligence tests correlate with achievement as well as traditional IQ? This question was examined by Naglieri (1999) who first reported that the correlations between achievement test scores with the CAS and KABC were as high as or higher than those found for the WISC-III and WJ-R. More recent findings are provided next.

Naglieri and Rojahn (2004) examined the relationships between the planning, attention, simultaneous, and successive (PASS) theory, as operationalized by the CAS, and achievement, as measured by the Woodcock-Johnson Tests of Achievement – Revised (WJ-R; Woodcock and Johnson 1989), using a nationally representative sample of 1,559 students. The correlation between the CAS Full Scale with the WJ-R Tests of Achievement was .71. More recently, Naglieri, Goldstein, DeLauder, and Schwebach (2006a) compared the Wechsler Intelligence Scale for Children – Third Edition (WISC-III; Wechsler 1991) to the CAS and the WJ-III Test of Achievement. The CAS Full Scale score correlated .83 with the WJ-III achievement scores compared to a coefficient of .63 for the WISC-III Full Scale IQ. The results suggest that when the same children took the two ability tests and those scores were correlated with the same achievement scores, both showed a strong relationship between ability and achievement, but the CAS correlated significantly higher (Naglieri et al. 2006a). Most recently, Naglieri et al. (2014) reported an average correlation between the CAS2 and achievement of .70.

The KABC-II Mental Processing Index (MPI) which excludes measures of knowledge correlated, on average, .68 with total achievement on the Peabody Individual Achievement Test – Revised (PIAT-R; Markwardt 1997), .70 with total achievement on the WJ-III Achievement Scale, .74 with total achievement on the Wechsler Individual Achievement Test – Second Edition (WIAT-II; Wechsler 2005), and .74 with total achievement on the Kaufman Test of Educational Achievement – Second Edition Comprehensive Form (KTEA-II;

Kaufman and Kaufman 2004b) (Kaufman and Kaufman 2004a, Tables 8-23 to 8-30; Kaufman and Kaufman 2004b, Table 7-25). Taken as a whole, the average correlation is .69.

The studies of the CAS, CAS2, KABC, and KABC-II summarized here illustrate that a cognitive approach to understanding children's intelligence is strongly correlated with achievement test scores. Interestingly, these studies show that cognitive processes are as effective for prediction of academic performance as traditional IQ tests even though the CAS, CAS2, KABC, and KABC-II do *not* include academically laden measures such as vocabulary and arithmetic. This provides an advantage for understanding achievement strengths and weaknesses for children who come from disadvantaged environments as well as those who have had a history of academic failure.

Are There Advantages to Second-Generation Intelligence Tests?

Having shown that *second-generation* ability tests correlate as well with achievement test scores as traditional IQ test which contain academic content, the next question to consider is do these tests have other advantages? For example, do these tests yield ability profiles that help understand the role a cognitive weakness may play in academic failure. A second important issue is related to fair assessment of diverse populations. More specifically, how do race differences on traditional IQ compare to those found for *second-generation* ability tests? And finally, can second-generation ability tests inform instruction and academic intervention? Each of these issues will be addressed next.

Do First- and Second-Generation Tests Detect Cognitive Problems That Underlie Academic Failure?

All intelligence tests give a full-scale score which is comprised of scales which in turn are comprised of subtests. The analysis of subtest and scale variation on tests such as the Wechsler Scales is a method called profile analysis that has been advocated by Kaufman (1994) and others (e.g., Sattler 1988) as a way to identify intellectual strengths and/or weaknesses. Information about strengths and weaknesses is then used to generate hypotheses that are integrated with other information so that decisions can be made regarding eligibility, diagnosis, and treatment. Despite the widespread use of this method, some have argued that subtest profile analysis does not provide useful information beyond that which is obtained from the IQ scores (e.g., McDermott et al. 1990; Dombrowski and Watkins 2013). Naglieri (1999) proposed that subtest analysis is problematic because of limitations in subtest reliability and validity and further suggested that what is needed is profile analysis based upon a sound theory of cognitive abilities, rather than individual subtest level analysis. Theoretically derived scales could be helpful if the ability test shows a specific pattern for a specific group of exceptional students which in turn could have implications for understanding the cognitive characteristics of the group, allow for guidance during the eligibility process (see Naglieri 2011), and guide interventions (Naglieri and Pickering 2010).

Recently, Naglieri (2011) summarized reports found in the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Wechsler 2003) technical manual, the Woodcock-Johnson III Tests of Cognitive Abilities (WJ-III; Woodcock et al. 2001) from Wendling et al. (2009), and CAS data from the technical manual and Naglieri, Otero, DeLauder, and Matto (2007). In the current chapter findings for students with autism spectrum disorders (ASD) from the WISC-IV (Wechsler 2003), the CAS (from Naglieri and Otero 2011), the WJ-III (from Wendling et al. 2009), and the KABC-II (technical manual) were added. The findings (see Fig. 20.1) must be considered with recognition that the samples were not matched on demographic variables across the various studies, the accuracy of the diagnosis may not have been verified, and some of the sample sizes were small. Notwithstanding these limitations, the findings provide important insights into the extent to which these various tests are likely to yield scale-level profiles that are distinctive, theoretically logical, and relevant to instruction.

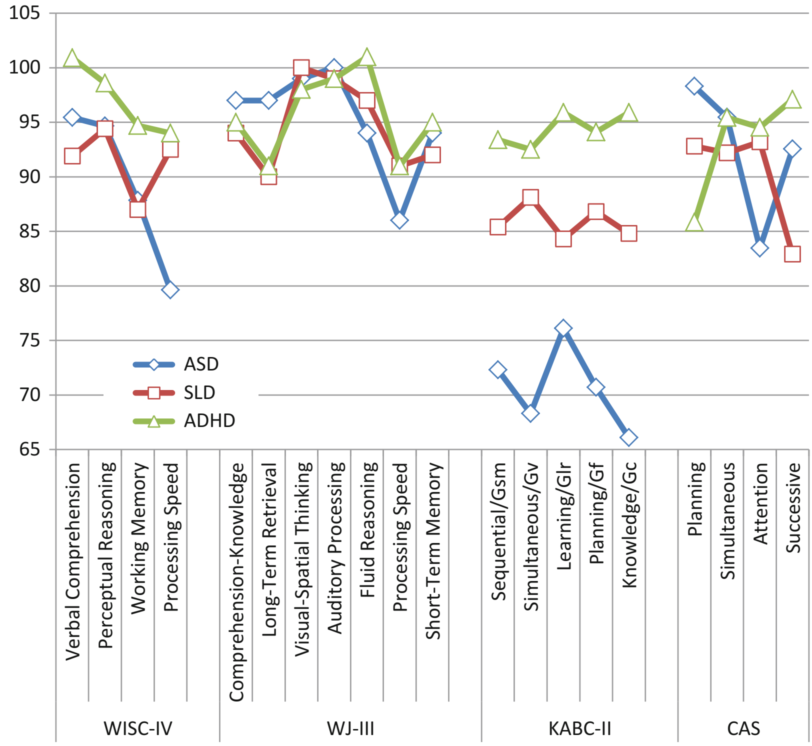


Fig. 20.1 Scale profiles on several measures of cognitive ability for students with SLD, ADHD, and autism

The results of the summary of scale profiles for the WISC-IV, WJ-III, KABC-II, and CAS provided in Fig. 20.2 suggest that some of these tests yield profiles that are more distinct than others across the three groups of exceptional children. The scores across all scales on the WJ-III for students with specific reading decoding difficulty (SLD) were all within the average range, and all of the KABC-II scores were in the 80s. The WISC-IV profile was lowest for the Working Memory Scale. The CAS profile showed variability across the four PASS scales with a very low score of 82.9 on the successive scale. These findings are consistent with the view that students with specific reading decoding failure also have considerable difficulty with tasks that involve sequencing of information (Das et al. 2007).

The intelligence test profiles for students with ADHD showed that all the scores for the scales on the WISC-IV, WJ-III, and KABC-II were with the average range. None of these tests provided evidence of a cognitive problem related to ADHD, except for a low score on the planning scale of the CAS. Difficulty with planning (e.g., executive function) for children with ADHD is consistent with Barkley's view that ADHD is a failure of self-control (Barkley 1997) which has been described as frontal lobe functioning (Goldberg 2009).

The results for individuals with autism spectrum disorders (ASD) show that processing speed scores on both the WISC-IV and the WJ-III were very low for individuals with ASD. This is similar to the findings for these two tests for individuals with ADHD. The low processing speed scores provide little insight into the cognitive characteristics of students with ASD and ADHD. Importantly, the low attention score on the CAS is consistent with the conceptualization that individuals with ASD have been described as having “difficulties in disengaging and shifting attention” (p. 214) (Klinger et al. 2009).

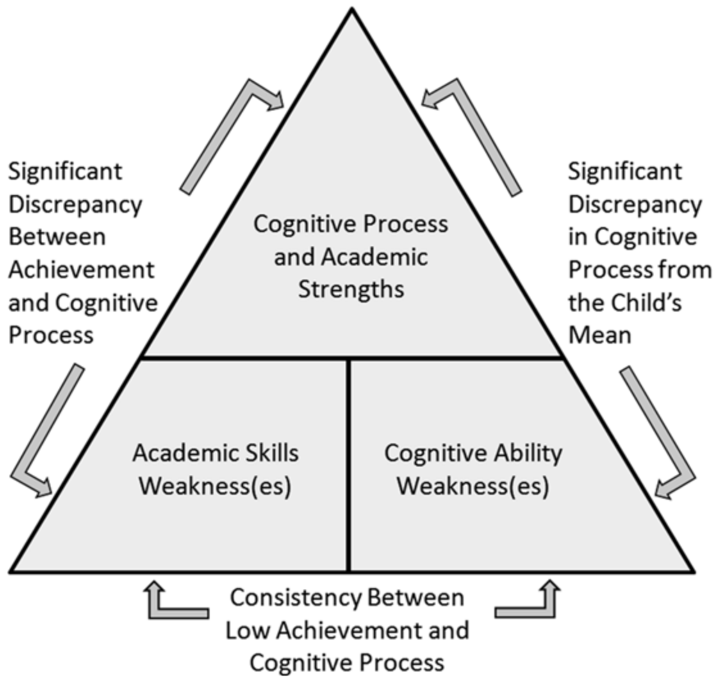


Fig. 20.2 Discrepancy/consistency model for specific learning disabilities

The findings for those with ASD, like the results for those with SLD and ADHD, show that the only test that had distinctive profiles for these different groups was the CAS.

Are Race Differences for Second-Generation Tests the Same as for Traditional IQ?

The need for intelligence tests that are appropriate for diverse populations of children has become progressively more important as the characteristics of the US population have changed, and recent Federal law (e.g., IDEA 2004) stipulates that assessments must be selected and administered so as to be nondiscriminatory on a racial or cultural basis. It is, therefore, critical that any measures used for evaluation be evaluated for test bias. This should include internal evidence such as reliability, item difficulty, and factor structure (see Jensen 1980) as well as mean score differences.

Some researchers have suggested that intelligence conceptualized on the basis of neuropsychological abilities is more appropriate for diverse populations (Fagan 2000; Naglieri and Otero 2011). Fagan (2000) and Suzuki and Valencia (1997) argued that measures of cognitive processes which do not rely on tests with language and quantitative content are more appropriate for assessment of culturally and linguistically diverse populations. Although there is considerable evidence for the validity of general intelligence as measured by traditional IQ tests (see Jensen 1980), researchers have traditionally found a mean difference of about 12–15 points between African-Americans and Whites on measures of IQ that include verbal, quantitative, and nonverbal tests (Kaufman and Lichtenberger 1999). Results for second-generation intelligence tests have been different.

The first evidence of smaller race differences for second-generation ability test was reported in the original KABC Manual. For children aged 2.5–12.5, without controlling for background variables, Whites (N =1,569) scored 7 points higher than African-Americans (N =807) and 3 points higher than Hispanics (N=160) on the global measure of mental processing (i.e., the total test score). These differences are considerably smaller than the differences of 16 points and 11 points, respectively, reported for the WISC-R Full Scale IQ (Kaufman and Kaufman 1983, Tables 4.36 and 4.37; Kaufman et al. 2005, Table 6.7). Similar findings were reported by Naglieri (1986) in a study of 172 fifth-grade students (86 whites and 86 blacks matched on basic demographic variables) who were administered the KABC and the WISC-R. The difference between the groups on WISC-R Full Scale was 9.1 but the difference for the KABC was 6.0. Results for the KABC-II (Kaufman and Kaufman 2004) showed a similar reduction in race/ethnic differences. When controlling for gender and mother's education, African-American children at ages 3–18 years earned mean MPIs that were only 5 points lower than the means for White children (A. S. Kaufman and Kaufman 2004a, Tables 8.7 and 8.8; A. S. Kaufman et al. 2005, Table 6.7). Similar findings have been reported for the CAS.

Naglieri, Rojahn, Matto, and Aquilino (2005) compared PASS scores on the CAS for 298 African-American children and 1,691 White children. Controlling for key demographic variables, regression analyses showed a CAS Full Scale mean score difference of 4.8 points in favor of White children. Similarly, Naglieri, Rojahn, and Matto (2007) examined the utility of the PASS theory with Hispanic children by comparing performance on the CAS of Hispanic and White children. The study showed that the two groups differed by 6.1 points using unmatched samples, 5.1 with samples matched on basic demographic variables, and 4.8 points when demographics differences were statistically controlled. Naglieri, Otero, DeLauder, and Matto (2007) compared scores obtained on the CAS when administered in English and Spanish to bilingual children ($N=40$) referred for reading difficulties. They found a 3.0-point difference between the CAS Full Scale scores and these scores were highly correlated (.96). Otero, Gonzales, and Naglieri (2012) replicated that study with another group of students referred for reading problems and found CAS Full Scale scores that differed by less than one point and a high correlation between the scores (.94). Results for the CAS2 Full Scale scores were reported in the test manual (Naglieri et al. 2014). For children and adolescents aged 5–18 years without controlling for demographic variables, African-Americans and non-African-Americans differed by 6.3 standard scores, and with controls for demographic characteristics, the difference was 4.5. Similarly, without controlling for demographic differences, Hispanics and non-Hispanics differed on the CAS Full Scale scores by 4.5 points, and with controls for demographic characteristics, the difference was 1.8.

Table 20.1 Mean score differences in standard scores by race on traditional IQ and second-generation intelligence tests

Test	Difference
Traditional	
SB-IV (matched)	12.6
WISC-IV (normative sample)	11.5
WJ-III (normative sample)	10.9
WISC-IV (matched)	10.0
Second generation	
KABC (normative sample)	7.0
KABC (matched)	6.1
KABC-2 (matched)	5.0
CAS2 (normative sample)	6.3
CAS (demographic controls)	4.8
CAS2 (demographic controls)	4.3

Notes: Stanford-Binet IV (SB-IV) from Wasserman (2000); (Woodcock-Johnson III) WJ-III from Edwards and Oakland (2006); Kaufman Assessment Battery for Children (KABC) matched from Naglieri (1986); Kaufman Assessment Battery for Children – 2 from (Lichtenberger et al. 2009); CAS from Naglieri, Rojahn, Matto, and Aquilino (2005); Wechsler Intelligence Scale for Children – IV (WISC-IV) from O'Donnell (2009)

The importance of the findings presented above for the CAS and KABC is best understood within the context of differences found on traditional intelligence tests. Table 20.1 provides a summary of standard score differences by race for the Stanford-Binet IV (SB-IV; Roid 2003), Woodcock-Johnson Tests of Cognitive Abilities – Third Edition (WJ-III; Woodcock et al. 2001), the WISC-IV (Wechsler 2003), the KABC and KABC-II (Kaufman and Kaufman 1983, 2004), and the CAS (Naglieri and Das 1997) and CAS2

(Naglieri et al. 2014). The results for the WISC-IV are reported by O'Donnell (1009), for the SB-IV by Wasserman (2000), and the WJ-III results are from Edwards and Oakland (2006). The race differences for the KABC normative sample were reported in that test's manual (Kaufman and Kaufman 1983), and the findings for the KABC-II were summarized by Lichenberger, Sotelo-Dynega, and Kaufman (2009). Differences for the CAS were reported by Naglieri, Rojahn, Matto, and Aquilino (2005) and in the test manual for the CAS2 by Naglieri, Das, and Goldstein (2014).

The results of research on race differences for the KABC-II and CAS illustrate that second-generation tests, in contrast to traditional IQ tests, provide a more equitable way to assess diverse populations of children. The findings suggest that as a group, traditional IQ tests showed differences in ability scores between the races that are about twice as large as that found for second-generation tests. All of the traditional tests included in this table have verbal/quantitative/nonverbal content, and two of these three types of questions demand knowledge that is very similar to that required by standardized achievement tests (see Naglieri and Bornstein 2003). It is reasonable to conclude that the approach to conceptualizing and measuring intelligence taken by the authors of these second-generation ability tests resulted in smaller race difference without a loss of prediction to achievement or in the case of the CAS sensitivity to learning problems, both of which are critical components of validity.

Do Second-Generation Tests Have Relevance to Academic Intervention?

One important purpose of assessment of ability is to help decide how a student learns best and what obstacles to learning may exist. Knowing the cognitive profile of an individual student should inform instruction. This means that in addition to teaching knowledge and skills, tailored instruction should help children, for example, "to plan and control, to think and inquire, to evaluate and reflect" (Scheid 1993, p. 3). This kind of approach teaches children knowledge and skills as well as effective ways of using the abilities a student has and managing any limitations in ability. The student is seen as an active participant who interprets information that is received, relates it to previously acquired facts, organizes and stores it for later use, develops ways of doing things, and critically examines information. Because this approach puts emphasis on both the academic skills the child must learn as well as the cognitive abilities the child uses in the act of learning, knowing the cognitive ability profile of a student is a critical element in a complex process that leads to effective teaching and learning (Naglieri 1999). The relationships between cognitive abilities as measured by the CAS and academic instruction have been reported in a series of research papers.

There are several resources for applying the PASS theory to academic instruction and remediation. The PASS Remedial Program (PREP; Das 1999) is an option as is the planning strategy instruction, also known as the planning facilitation method, described by Naglieri and Pickering (2003). Other resources include, for example, Kirby and Williams' (1991) Learning Problems: A Cognitive Approach, Cognitive Strategy Instruction That Really Improves Children's Academic Performance – Second Edition by Pressley and Woloshyn (1995), *Helping Students Become Strategic Learners* (Scheid 1993), and Naglieri and Pickering's (2010) book *Helping Children Learn: Intervention Handouts for Use in School and Home*. The first two methods are based on empirical studies, while the remaining books contain cognitive approaches to academic interventions. The methods use structured and directed instructions based on PREP or structured but not scripted planning strategy instruction. In order to provide more details and research underlying strategy instruction and PREP, these two methods will be discussed in more detail.

Strategy Instruction. The connection between planning from PASS and interventions to improve the use of strategies has been examined in a series of studies. These investigations have involved both math and reading comprehension and have focused on the concept that children can be taught to be more strategic when they complete academic tasks and that the facilitation of plans positively impacts academic performance. The essential concept was based on the idea that *teaching* students to use strategies directly was not as advantageous as *encouraging* students to approach their work strategically. The method is designed so that children discover the value of strategy use without being specifically instructed to do so. The students are encouraged to examine the demands of the task in a strategic and organized manner. Research on this intervention method and its relationship to PASS scores on the CAS has been carefully examined in a number of important research studies.

The first two studies using planning strategy instruction showed that children's performance in math calculation improved substantially (Naglieri and Gottling 1995, 1997). The children in these two studies attended a special school for those with learning disabilities. Students completed mathematics worksheets in sessions over about a two-month period. The method designed to teach planning was applied in individual 1 on 1 tutoring sessions (Naglieri and Gottling 1995) or in the classroom by the teacher (Naglieri and Gottling 1997) two to three times per week in half hour blocks of time. Students were encouraged to recognize the need to plan and use strategies when completing mathematic problems during the intervention periods. The teachers provided probes that facilitated discussion and encouraged the students to consider various ways to be more successful. More details about the method are provided by Naglieri and Gottling (1995, 1997) and by Naglieri and Pickering (2010).

The relationship between strategy instruction and the PASS profiles for children with learning disabilities and mild mental impairments was also studied by Naglieri and Johnson (2000). The purpose of their study was to determine if children with cognitive weaknesses in each of the four PASS processes, and children with no cognitive weaknesses, showed different rates of improvement in math when given the same group planning strategy instruction. The findings from this study showed that children with a cognitive weakness in planning improved considerably over baseline rates, while those with no cognitive weakness improved only marginally. Similarly, children with cognitive weaknesses in simultaneous, successive, and attention showed substantially lower rates of improvement. The importance of this study was that the five groups of children responded very differently to the same intervention. Stated another way, the PASS processing scores were predictive of the children's response to this math intervention (Naglieri and Johnson 2000).

The effects of planning strategy instruction on reading comprehension were reported by Haddad, Garcia, Naglieri, Grimditch, McAndrews, and Eubanks (2003). This study assessed whether an instruction designed to facilitate planning would have differential benefits on reading comprehension and if improvement was related to the PASS processing scores of each child. The researchers used a sample of general education children sorted into groups based on their PASS scale profiles using the CAS. Even though the groups did not differ by CAS Full Scale scores or pretest reading comprehension scores, children with a planning weakness benefited substantially (effect size of 1.4) from the instruction designed to encourage the use of strategies and plans. In contrast, children with no PASS weakness or a successive weakness did not benefit as much (effect sizes of .52 and .06, respectively). These results further support previous research suggesting that the PASS profiles are relevant to instruction.

Iseman and Naglieri (2011) examined the effectiveness of the strategy instruction for students with LD and ADHD randomly assigned to an experimental group or a control group which received standard math instruction. They found large pre-post effect sizes for students in the experimental group (0.85), but not the control group (0.26) on classroom math worksheets, as well as standardized test score differences in Math Fluency (1.17 and .09, respectively) and Numerical Operations (.40 and −.14, respectively). One year later the experimental group continued to outperform the control group. These findings strongly suggested that students with LD and ADHD in the experimental group evidenced greater improvement in math worksheets, far transfer to standardized tests of math, and at follow-up 1 year later than the control group. The findings also illustrate the effectiveness of strategy instruction especially for those with low planning scores on the CAS.

The results of these planning strategy instruction studies using academic tasks suggest that changing the way aptitude is conceptualized (e.g., as the PASS rather than traditional IQ) and measured (using the CAS) increases the probability that an aptitude-by-treatment interaction (ATIs) is detected. Past ATI research suffered from conceptualizations of aptitudes based on the general intelligence model which did not adequately differentiate cognitive abilities which *are* related to instruction. The traditional IQ approach is very different from the PASS theory as measured by the CAS. The summary of studies provided here are particularly different from previous ATI research that found students with low general ability improve little, whereas those with high general ability respond more to instruction. In contrast, children with a weakness in one of the PASS processes (planning) benefited *more* from instruction compared to children who had no weakness or a weakness in a different PASS process. The results of these studies also suggest that the PASS profiles can help predict which children will respond to the academic instruction and which will not. This offers an important opportunity for researchers and practitioners interested in the design of instruction as suggested by Naglieri and Pickering (2003).

PREP. PREP was developed as a cognitive remedial program based on the PASS theory (Das et al. 1994). These researchers summarized research which showed that students could be trained to use successive and simultaneous processes more efficiently, which resulted in an improvement in their performance on that process and transferred to specific reading tasks. PREP aims to improve the use of cognitive processing strategies (e.g., simultaneous and successive processes) that underlie reading. The tasks in the program teach children to focus their attention on the sequential nature of many tasks, including reading which helps the children better utilize successive processing – a very important cognitive process needed in reading decoding. PREP is also founded on the premise that the transfer of principles is best facilitated through inductive, rather than deductive, inference. The program is, therefore, structured so that tacitly acquired strategies are likely to be used in appropriate ways. For example, the tasks teach children to focus on the sequences of information included in a variety of tasks, including reading.

Support for PREP summarized elsewhere (Naglieri 2011) has shown the effectiveness of the instructional method for children with reading decoding problems. Children who received PREP in comparison to a regular reading program improved significantly on Word Attack and Word Identification tests from the Woodcock Reading Mastery Test – Revised. Learning disabled children who were randomly assigned to a PREP training and a control group that received regular classroom instruction showed significant improvement in reading decoding of real and pseudowords. When PREP was compared to a meaning-based reading program using two carefully matched groups of first-grade children, the results showed a significant improvement in reading scores for the PREP group and the gain in reading was greater than it was for the meaning-based control group. Specific relevance to the children's CAS profiles was also demonstrated by the fact that those children with a higher level of successive processing at the beginning of the program benefited the most from the PREP instruction, but those with the most improvement in the meaning-based program had higher levels of planning. Taken as a whole, these studies support the effectiveness of PREP in remediating deficient reading skills during the elementary school years and the connection between the PASS theory and intervention.

Do Second-Generation Ability Tests Aid in Determination of a Specific Learning Disability?

One of the greatest challenges to traditional IQ has been the identification of a specific learning disability, defined in IDEA 2004 as follows:

> *Specific learning disability* means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Such term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such term does not include a learning problem that is primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

Perhaps the most essential problem with using a traditional IQ test for this purpose is the fact that alignment of the test results to the Federal regulations has been quite difficult. This is especially true because specific learning disability is defined as a disorder in basic psychology processes, but traditional IQ tests were not developed to measure basic psychological processes. Even though it is clear that the identification of a disorder in one or more of the basic psychological processes is essential for SLD eligibility determination (Fuchs and Young 2006; Hale et al. 2006), practitioners have been constrained by the content of traditional IQ tests which do not align with the definition.

It is logical that assessment of basic psychological processes as stipulated in IDEA 2004 be a part of any comprehensive assessment designed to determine if a child has a specific learning disability. Moreover, IDEA 2004 requires the use of a variety of assessment tools that are technically sound and nondiscriminatory to gather functional, developmental, and academic information when special education eligibility is being determined. This information should be integrated with other important findings about the child to ensure that a comprehensive evaluation is obtained. In short, documentation of a basic psychological processing disorder and academic failure is essential for SLD determination (Hale et al. 2006; Naglieri 1999, 2011), and this can be best accomplished with second-generation ability tests.

Essential to the description of a specific learning disability (SLD) is the presence of a pattern of strengths and weakness in basic psychological processes and academic skills. Of all the possible tools described in this chapter, the PASS theory as measured by the CAS2 is best suited because it yields profiles for students with SLD, works well with minority students, and has intervention implications (Haung et al. 2010; Naglieri and Otero 2011). There are three main components to eligibility determination based on this conceptualization initially presented by Naglieri (1999). First, the student has significant intraindividual *discrepancy* among the PASS scales, and the lowest PASS ability score is substantially below average. Second, there is a *discrepancy* between good PASS scores and weak achievement. Third, there is a *consistency* between poor PASS scores and academic deficits (Naglieri 1999, 2011) as illustrated in Fig. 20.2.

An intraindividual discrepancy is examined by comparing a student's four PASS scale standard scores to determine if there exists a cognitive weakness. The purpose of these analyses is to identify PASS cognitive processing strengths (scores that are significantly greater than the student's mean score and fall above the normative average) or weaknesses (scores that are significantly lower than the student's mean score and fall below the normative average). For example, consider a student has standard scores of 114 (planning), 116 (simultaneous), 109 (attention), and 94 (successive). The successive score is 14.25 standard score points below the child's mean of 108.25 which is significant but that score of 94 is within the average range. Academic achievement scores are similar to the successive score (low portion of the average range). This would *not* be considered evidence of a disorder in one or more basic psychological processes and academic failure. In contrast, a cognitive weakness is found when, for example, a student has standard scores of 102, 104, 97, and 82 for planning, simultaneous, attention, and successive, respectively, in which case the successive score is considered a weakness and there are comparable academic achievement test scores. The successive and academic weaknesses in contrast to planning, simultaneous, attention, and academic scores that are average or higher would suggest the existence of a disorder in one or more of the basic psychological processes with academic failure as described in IDEA.

The method described above is referred to as the discrepancy/consistency model (Naglieri 1999). This method is useful for the identification of specific learning disabilities because it ensures a systematic examination of variability of both cognitive and academic achievement test scores. Determining if the cognitive processing scores differ significantly is accomplished using the method originally proposed by Davis (1959) and Silverstein (1982), popularized by Kaufman (1979), and modified by Silverstein (1993). This so-called ipsative method determines when the child's scores are reliably different from the child's average score. This technique has been applied to a number of tests including, for example, the WISC-IV (Naglieri and Paolitto 2005), the CAS (Naglieri and Das 1997), and the SB5 (Roid 2003). It is important to note that in the discrepancy/consistency model described by Naglieri (1999), the ipsative approach is applied to the PASS scales which represent four neurocognitive PASS constructs, not the subtests. This changes the method from one that relies on a clinical interpretation of the meaning of subtest variability to analysis of scales that have been theoretically defined and have higher reliability and validity. This distinction is important because the criticisms of the ipsative method (McDermott et al. 1990) have centered around subtest, not scale-level analysis.

Naglieri (1999) and Flanagan and Kaufman (2004) stressed the importance of recognizing that because a low score relative to the child's mean could still be within the average range, adding the requirement that the weakness in a processing test score is also well below average is important. In a study of PASS profiles for the CAS standardization and validity samples, Naglieri (2000) found that those students who had a PASS weakness were likely to have significantly lower achievement scores and more likely to have been identified as exceptional. That study was described by Carroll (2000) as one which illustrated what a more successful profile method could be. Davison and Kuang (2000) suggested that "adding information about the absolute level of the lowest score improves identification over what can be achieved using ipsative profile pattern information alone (p. 462)." Importantly, when Huang et al. (2010) studied PASS profiles on the CAS for large samples of students in regular education ($N=1{,}692$) and those with specific learning disabilities ($N=367$), they found ten core PASS profiles for those in regular educational and eight unique profiles from students with SLD. Huang et al. concluded that "a student with a true LD has a relatively high chance of being accurately identified when using profiles analysis on composite [PASS] scores" (p. 28). They added that their "analysis has provided evidence for the use of the PASS theory and that it appears that it has sufficient applications for diagnosis for students suspected of having a LD" (p. 28).

Section Summary

The topics covered thus far provide evidence that second-generation ability tests should be considered viable methods of evaluating children and adolescents for three important reasons. First, the KABC and CAS correlate strongly with achievement even though they do not have academic content, which suggests they have excellent validity. Second, the CAS and KABC yield small differences between Black and White (CAS and KABC and their second editions) as well as Hispanic and White (CAS, KABC-II, and CAS2) groups which provides evidence that these measures are appropriate for non-biased assessment. Third, the evidence presented shows that CAS scores reveal the weakness children with specific learning disability in reading decoding have is different from that experienced by those with other types of SLD (Haung et al. 2010) as well as ADHD and autism (Fig. 20.3).

The remainder of this chapter will focus on one of the two second-generation ability tests. Although the KABC and CAS both provide substantial advantages beyond traditional IQ, only the CAS has demonstrated specific PASS profiles for students with disabilities—it yields the smallest differences by race/ethnicity, and there is a history of research showing the relevance of PASS scores to academic instruction. For these reasons the remainder of this chapter will provide

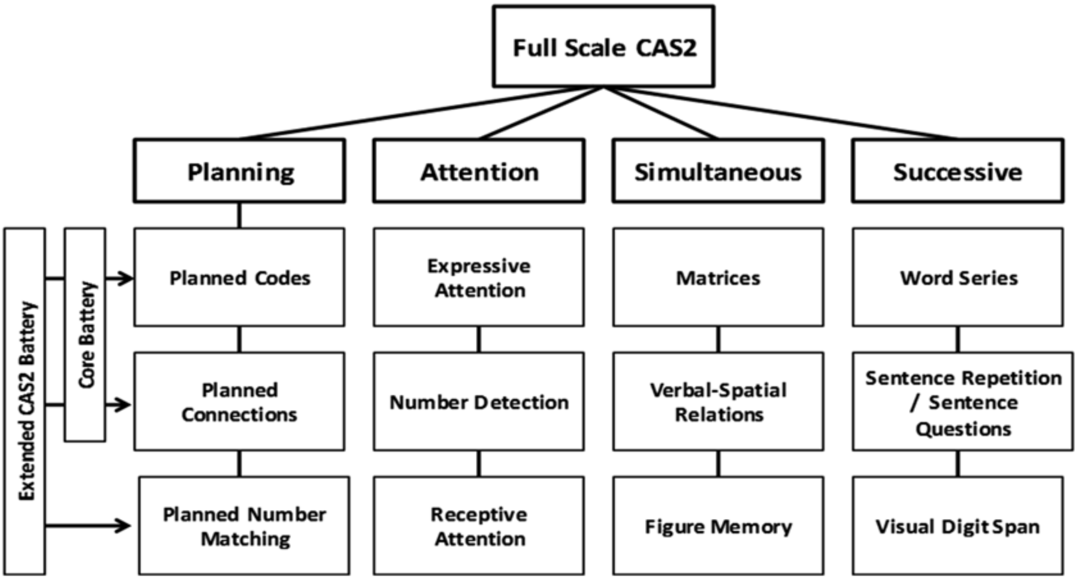


Fig. 20.3 Structure of the Cognitive Assessment System – Second Edition

a more detailed discussion of how the four PASS abilities can be measured by the CAS2.

Operationalization of the PASS Neurocognitive Abilities

The PASS theory was first operationalized by Naglieri and Das in 1997 with the publication of the CAS (and more recently the CAS2 (Naglieri et al. 2014)). Some of the research about the scores obtained from the tests developed to measure PASS was described earlier in this chapter. Although the PASS theory is more fully described elsewhere (Naglieri 1999; Naglieri and Das 1997; Naglieri and Otero 2011; Naglieri et al. 2014) and in Otero's chapter in this book, the remainder of this chapter includes a description of how each of the four PASS abilities are measured (see Fig. 20.3) in the CAS2 and how this relates to traditional IQ.

Planning

In PASS theory, planning is a neurocognitive ability used to determine and apply strategies to solve problems and self-monitor and self-correct as needed (Naglieri et al. 2014). This includes control of actions and thoughts so that efficient solutions to problems can be achieved. Planning provides the means to solve problems for which no method or solution is immediately apparent and may involve retrieval of information as well as utilization of the other PASS abilities to process the information. Planning ability is also important when individuals reflect on events, recognizing what worked, and what did not work, and considering better problem solving in the future. The frontal lobes of the brain are directly involved in planning ability (Naglieri and Otero 2011).

The essence of tasks that measure planning is that the student must solve novel problems for which there is no previously acquired strategy and there should be minimal constraints placed on the way the student completes the task. The score a planning test yields should reflect efficiency, measured by how a student went about completing the tests and how effective the solutions were. The following tasks are used in the CAS2 to evaluate planning ability:

Planned Codes. This subtest contains four items, each with its own set of codes and particular arrangements of rows and columns. A legend at the top of each page shows which letters correspond to which codes (e.g., A, B, C, D with OX, XX, OO, XO, respectively). Just below the legend are seven rows and eight columns of letters without the codes. Children write the corresponding codes in empty boxes beneath each of the letters. The items differ in the correspondence of letters to codes and the position of the letters on the page. Students have 60 s per item to complete as many empty code boxes as possible.

Planned Connections. The Planned Connections subtest requires the student to connect numbers in sequence that appears in a quasi-random order (e.g., 1–2–3, etc.). For the more complex items, the child connects numbers and letters in sequential order, alternating between numbers and letters (e.g., 1–A–2–B, etc.). Any errors made by the child are corrected as they progresses through the task. The items are constructed so that children never complete a sequence by crossing one line over the other. This provides a means of reducing the areas to be searched when looking for the next number or letter. The subtest score is based on the total amount of time used to complete the task.

Planned Number Matching. The student's task is to find and underline two numbers that are the same in each row. Each item is composed of eight rows of numbers, with six numbers per row. Two of the six numbers in each row are the same. The length of numbers differs on the various rows. Numbers increase in digit length from one digit on the first row of item 1 to seven digits on the eighth row of item 4. There are four rows for each digit length and a total of four pages of numbers. Children ages 5–7 are provided an example followed by two test items. Ages 8–18 are provided an example and two test items. Each row of numbers was carefully developed to maximize benefits of strategy usage in the identification of correct matches. This approach resulted in items with some rows that contain numbers that start with unique numbers, some rows that include numbers with similar digit strings, and some rows that contain numbers that end with similar numbers.

Attention

Attention is a neurocognitive ability used to selectively focus on a particular stimulus while inhibiting responses to competing stimuli presented over time (Naglieri et al. 2014). Attention is a basic component of intelligent behavior involving allocation of resources and effort. Arousal, attention, effort, and capacity are concepts that have a complex relationship and importance for understanding behavior. Luria stated that optimal conditions of arousal are needed before the more complex forms of attention involving "selective recognition of a particular stimulus and inhibition of responses to irrelevant stimuli" (Luria 1973, p. 271) can occur. Attention is conceptualized as a mental activity that provides focused, selective cognitive activity over time and resistance to distraction. The process is involved when a person must demonstrate focused, selective, sustained, and effortful activity. The longer the attention needed, the more the activity necessitates vigilance. Intentions and goals mandated by the planning process control attention, whereas knowledge and skills play an integral part in the process as well.

Tasks that measure attention include target and nontarget stimuli that are multidimensional with the requirement that the person has to identify one aspect of the target (e.g., the color blue) and resist responding to distractions (e.g., a word red written in blue ink) as in the Stoop test (Lezak 1995). This kind of a task requires selective focus of attention over time, an ability that is necessary for learning to take place. The following tasks are used in the CAS2 to evaluate attention:

Expressive Attention. The Expressive Attention subtest consists of two age-related sets of three items. Students ages 5–7 years are presented with three items consisting of seven rows that each contain six pictures of common animals, with each picture depicted as either big (1 in. by 1 in.) or small (1/2 in. by 1/2 in.). In each of three items, the student is required to identify whether the animal depicted is big or small in real life, ignoring the relative size of the picture on the page. In item 1, the pictures are all the same size.

In item 2, the pictures are sized appropriately (i.e., big animals are depicted with big pictures, and small animals are depicted with small pictures). In item 3, the realistic size of the animal often differs from its printed size. Students ages 8–18 years are presented with three items consisting of eight rows of five words each. In item 1, students are asked to read four black-and-white color words (blue, yellow, green, and red) that are presented in random order. In item 2, students are asked to name the colors of four colored rectangles (printed in blue, yellow, green, and red) that are presented in random order. In item 3, the four color words are printed in a different-color ink than the color-word name and are presented in random order. In this item, students are required to name the color of the ink in which the word is printed rather than read the word.

Number Detection. Each Number Detection item presents the student with a page of approximately 200 numbers. Students are required to underline specific numbers (ages 5–7 years) or specific numbers in a particular font (ages 8–18 years) on a page with many distractors. There are four pages of numbers, each of which is scored for the number correct, number of false detections, and time.

Receptive Attention. The Receptive Attention subtest consists of four item sets, each containing 60 picture pairs (ages 5–7 years) or 180 letter pairs (8–18 years). Both versions require the student to underline pairs of objects or letters that either are identical in appearance or are the same from a lexical perspective (i.e., they have the same name). There are four pages of numbers, each of which is scored for the number correct, number of false detections, and time.

Simultaneous

Simultaneous is a neurocognitive ability used to integrate separate stimuli into a single whole or interrelated group (Naglieri et al. 2014). The essence of simultaneous processing is that separate elements must be combined into a conceptual whole. This ability is involved in visual-spatial tasks as well as those language activities that require comprehensive of grammatical structures. The spatial aspect of simultaneous ability involves both the perception of stimuli as a group or whole and the formation of visual images. The grammatical dimension of simultaneous processing allows for the integration of words into ideas through the comprehension of word relationships, prepositions, and inflections, so the person can obtain meaning.

Tasks designed to measure simultaneous processing often have visual-spatial content. One well-known measure of simultaneous processing is progressive matrices. Traditional intelligence tests often include subtests that use the progressive matrix format, as do many nonverbal intelligence tests such as the Naglieri Nonverbal Ability Test (Naglieri 2011). These tests are often categorized as perceptual reasoning or nonverbal, but from PASS, matrices measure simultaneous ability. This ability can also be measured using verbal content which requires comprehension of the grammatical components of language such as comprehension of word relationships and understanding of prepositions and inflections (Naglieri 1999). The Verbal-Spatial Relations subtest on the CAS is an example of this type of a subtest (Naglieri et al. 2014). This arrangement of subtests allows for measurement of simultaneous ability across verbal and nonverbal contents. The tests used to evaluate simultaneous neurocognitive ability on the CAS2 are as follows:

Matrices. Matrices is a multiple-choice subtest that utilizes shapes and geometric elements that are interrelated through spatial or logical organization. Students are required to analyze the relationship among the parts of the item and solve for the missing part by choosing the best of five options. The raw score is the total number of items correctly answered.

Verbal-Spatial Relations. Verbal-Spatial Relations is a multiple-choice subtest in which each item consists of six drawings and a printed question at the bottom of each page. The examiner reads the question aloud, and the child is required to select the option that matches the verbal description. The items require the evaluation of logical grammatical relationships (e.g., "which picture shows a ball in a basket under a table?"), which demands simultaneous processing with verbal content. The raw score is the total number of items correctly answered.

Figure Memory. For each Figure Memory item, the examiner presents the student with a two- or three-dimensional geometric figure for 5 s. The picture is then removed, and the student is presented with a response page that contains the original figure embedded in a large, more complex geometric pattern. The student is required to trace the original figure with a red pencil in the Figure Memory Response Form. The raw score is the total number of items correctly answered.

Successive

Successive is a neurocognitive ability used to work with information that is arranged in a specific serial order where each part follows the other in a strictly defined order (Naglieri et al. 2014). Successive processing is involved in the perception of stimuli in sequence as well as the formation of sounds and movements into a specific order. This type of ability is necessary for the recall of information in order as well as phonological analysis and the syntax of language (Das et al. 1994). Deficits with successive processing are also associated with early reading problems in young children, as it requires a child to learn sounds in a sequential order.

Tasks used to measure successive processing include digit span forward (as well as the recall of numbers, words, or hand movements) which is found on many tests of ability. These tests are sometimes described as measures of working memory or sequential processing (a concept very close to successive processing in PASS theory). Sometimes a backwards version is included which involves successive as well as planning processing abilities (Schofield and Ashman 1987). The successive tasks included in the CAS and CAS2 provide a way to measure this ability using tests that demand repeating a sentence using the correct series of words (Sentence Repetition) as well as comprehension of sentences that are understood only by appreciating the sequence of words (Sentence Questions). Additionally, CAS2 has a visual digit span test, allowing for measurement of successive processing across auditory and visual modalities. The tests used in the CAS2 to measure successive processing ability are as follows:

Word Series. The Word Series subtest utilizes nine single-syllable, high-frequency words: book, car, cow, dog, girl, key, man, shoe, and wall. The examiner reads aloud a series of two to nine of these words at the rate of one word per second. The student is required to repeat the words in the same order as stated by the examiner. The raw score is the total number of items correctly answered.

Sentence Repetition. The Sentence Repetition subtest (administered only to ages 5–7 years) requires the student to repeat syntactically correct sentences containing little meaning, such as "The blue is yellowing." The raw score is the total number of items correctly answered.

Sentence Questions. The Sentence Questions subtest (administered only to ages 8–18 years) requires the student to listen to sentences that are syntactically correct but contain little meaning and answer questions about the sentences. For example, the student is read the sentence "The blue is yellowing" and then asked the following question: "Who is yellowing?" The raw score is the total number of items correctly answered.

Visual Digit Span. Visual Digit Span subtest requires the student to recall a series of numbers in the order in which they were shown using the Stimulus Book. Each item that is 2–5 digits in length is exposed for the same number of seconds as there are digits. Items with six digits or more are all exposed for a maximum of 5 s. The raw score is the total number of items correctly answered.

The CAS2 subtests described above can be combined into an 8-subtest Core Battery or a

12-subtest Extended Battery to yield four scores following the PASS theory, planning, attention, simultaneous, and successive, and a total score called the Full Scale. The subtests are all individually administered tests designed explicitly to yield scores to evaluate the four PASS neurocognitive abilities for children and adolescents aged 5 years 0 months through 18 years 11 months. It was normed on a representative sample of 1,342 students. The test manual provides a complete summary of reliability and validity of the CAS2 as well as interpretive and intervention information. See Naglieri et al. (2014) for more details.

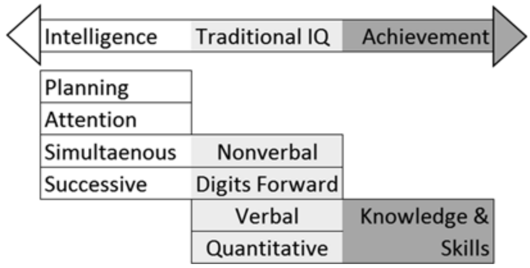


Fig. 20.4 Intelligence achievement continuum

Closing Thoughts

The purpose of this chapter was to organize 100 years of progress in the area of IQ tests, twentieth-century traditional ideas about intelligence, and a second generation of intelligence tests. The essence of the discussion has been about the tools and concepts we have used in this most important field of applied psychology. I have argued that there are several important issues that need to be recognized and will be reiterated here.

First, traditional IQ began July 20, 1917, with the development of the verbal (quantitative) and nonverbal IQ test format. This format has dominated the IQ testing industry since that time and has been used in all individual- and group-administered IQ tests.

Second, the tests developed for the US Army were designed to test many recruits in the shortest amount of time and with the least amount of effort needed for scoring. There was no theory of intelligence that guided the selection or development the Army Alpha and Beta tests. These tests have been accepted as measures of intelligence and in fact the IQ score has become synonymous with the term intelligence.

Third, traditional IQ tests include questions that are very similar to tests found in achievement tests, especially, for example, vocabulary, word analogies, and math word problems. The role of knowledge needed to answer these types of questions was recognized as undesirable by the original authors of the Army Alpha. Despite the fact that such tests are often indistinguishable from measures of *achievement*, users of traditional IQ tests have ignored this problem and compounded the issue by calling such tests measures of *verbal intelligence*.

Fourth, about 80 years after the birth of traditional IQ, a second generation of intelligence tests appeared. The first was the KABC (Kaufman and Kaufman 1983) and second was the CAS (Naglieri and Das 1997). These tests were designed with a conceptualization of intelligence (KABC) or a specific theory of intelligence (CAS); and importantly, they did not include the verbal and arithmetic test items found in traditional IQ. On a continuum from pure intelligence to pure achievement, the second-generation tests were clearly distinct from tests of achievement and traditional IQ (see Fig. 20.4). Traditional IQ tests share some overlap with second-generation IQ tests (mainly the nonverbal portion of traditional IQ tests and simultaneous scales of the CAS and KABC). The verbal and quantitative portions of traditional IQ share overlap with achievement tests. The academic content of traditional IQ tests poses considerable problem for test validity and assessment of diverse populations as well as those with learning difficulties.

Fifth, research has clearly supported second-generation intelligence tests over traditional IQ. Newer tests offer several advantages including (a) a theory that can be used to create scales on a test that represent a psychological construct, (b) greater fairness to minorities and to those with limited academic skills, (c) scores that represent different abilities according to the theory upon which the test was developed, (d) greater ability to identify special populations of individuals with intellectual disorders related to behavioral or academic disabilities, and (e) enhanced ability to link second-generation intelligence test scores to interventions.

The challenge faced by second-generation intelligence tests, despite their clear advantages over traditional IQ, is inertia. Traditional IQ has 100 years of use and acceptance and countless numbers of research studies and books written about them and their interpretation. But as Neil deGrasse Tyson, author of the new guide to the Cosmos, recently commented (2014) on the value of traditional wisdom "In practically every idea we have as humans, the older version of it is *not* better than the new version" (p. 80). It is time for the field of intelligence testing to embrace new ideas of what intelligence may be and how best to measure it.

References

Barkley, R. A. (1997). *ADHD and the nature of self-control*. New York: Guilford.

Boake, C. (2002). From the Binet-Simon to the Wechsler-Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology, 24*, 383–405.

Carroll, J. B. (2000). Commentary on profile analysis. *School Psychology Quarterly, 15*, 449–456.

Das, J. P. (1999). *PASS reading enhancement program*. Deal: Sarka Educational Resources.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes*. Boston: Allyn & Bacon.

Das, J. P., Janzen, T., & Georgiou, G. K. (2007). Correlates of Canadian native children's reading performance: From cognitive styles to cognitive processes. *Journal of School Psychology, 45*, 589–602.

Davis, F. B. (1959). Interpretation of differences among averages and individual test scores. *Journal of Educational Psychology, 50*, 162–170.

Davison, M. L., & Kuang, H. (2000). Profile patterns: Research and professional interpretation. *School Psychology Quarterly, 15*, 457–464.

Dombrowski, S. C., & Watkins, M. W. (2013). Exploratory and higher order factor analysis of the WJ-III full test battery: A school-aged analysis. *Psychological Assessment 2013, 25*(2), 442–455.

Edwards, O. W., & Oakland, T. D. (2006). Factorial invariance of Woodcock-Johnson III scores for African Americans and Caucasians Americans. *Journal of Psychoeducational Assessment, 24*, 358–366.

Fagan, J. R. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law, 6*, 168–179.

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. Hoboken: Wiley.

Fuchs, D., & Young, C. L. (2006). On the irrelevance of intelligence in predicting responsiveness to reading instruction. *Exceptional Children, 73*, 8–30.

Goldberg, E. (2009). *The new executive brain: Frontal lobes in a complex world*. New York: Oxford University Press.

Haddad, F. A., Garcia, Y. E., Naglieri, J. A., Grimditch, M., McAndrews, A., & Eubanks, J. (2003). Planning facilitation and reading comprehension: Instructional relevance of the PASS theory. *Journal of Psychoeducational Assessment, 21*, 282–289.

Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools, 43*(7), 753–770.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American Children*. Baltimore: Brookes.

Haung, L. V., Bardos, A. N., & D'Amato, R. C. (2010). Identifying students with learning disabilities: Composite profile analysis using the cognitive assessment system. *Journal of Psychoeducational Assessment, 28*, 19–30.

IDEA. (2004). Individuals with Disabilities Education Improvement Act of 2004 (P.L. 108–446).

Iseman, J., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD: A randomized controlled study. *Journal of Learning Disabilities, 44*, 184–195.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Circle Pines: American Guidance.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.

Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman assessment battery for children second edition. Circle Pines, MN: American Guidance.

Kaufman, A. S., & Kaufman, N. L. (2004a). *Kaufman assessment battery for children* (2nd ed.). Circle Pines: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman test of educational achievement comprehensive form* (2nd ed.). Circle Pines: American Guidance Service.

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. New York: Wiley.

Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II assessment*. New York: Wiley.

Kirby, J. R., & Williams, N. H. (1991). *Learning problems: A cognitive approach*. Toronto: Kagan and Woo.

Klinger, L., O'Kelly, S. E., & Mussey, J. L. (2009). Assessment of intellectual functioning in Autism spectrum disorders. In S. Goldstein, J. A. Naglieri, &

S. Ozonoff (Eds.), *The assessment of autism spectrum disorders* (pp. 209–252). New York: Guilford.

Lichtenberger, E. O., Sotelo-Dynega, M., & Kaufman, A. S. (2009). The Kaufman assessment battery for children – second edition. In J. *A. Naglieri & S. Goldstein (Eds.), A practitioner's guide to assessment of intelligence and achievement*. New York: Wiley.

Luria, A. R. (1963). *Restoration of function after brain injury*. New York: Pergamon Press.

Luria, A. R. (1966/1980). *Higher cortical functions in man* (2nd ed., Revised and Expanded). New York: Basic Books.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books.

Markwardt, F. C. (1997). *Peabody individual achievement test-revised (PIAT-R)*. San Antonio: The Psychological Corporation.

Matarazzo, J. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: Williams & Wilkins.

McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment, 8*(3), 290–302.

Naglieri, J. A. (1986). WISC-R and K-ABC comparison for matched samples of Black and White children. *Journal of School Psychology, 24*, 81–88.

Naglieri, J. A. (1999). *Essentials of CAS assessment*. New York: Wiley.

Naglieri, J. A. (2000). Can profile analysis of ability test scores work? An illustration using the PASS theory and CAS with an unselected cohort. *School Psychology Quarterly, 15*, 419–433.

Naglieri, J. A. (2003). Fair assessment of gifted minority children using the NNAT. *Gifted Education Press Quarterly, 17*, 2–7.

Naglieri, J. A., & Conway, C. (2009). The cognitive assessment system. In J. A. Naglieri & S. Goldstein (Eds.), *A practioner's guide to assessment of intelligence and achievement* (pp. 3–10). New York: Wiley.

Naglieri, J. A. (2011). The discrepancy/consistency approach to SLD identification using the PASS theory. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 145–172). Hoboken: Wiley.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21*, 244–260.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive assessment system*. Itasca: Riverside Publishing Company.

Naglieri, J. A., & Das, J. P. (2001). The Das-Naglieri Cognitive Assessment System in theory and practice. In J. Andrews (Ed.), *Ability, achievement, and behavior assessment: A practical handbook* (pp. 36–68).

Naglieri, J. A., & Goldstein, S. (2011). Assessment of cognitive and neuropsychological processes. In S. Goldstein, J. A. Naglieri, & M. DeVries (Eds.), *Understanding and managing learning disabilities and ADHD in adolescence and adulthood* (2nd ed., pp. 137–160). New York: Wiley.

Naglieri, J. A., & Gottling, S. H. (1995). A cognitive education approach to math instruction for the learning disabled: An individual study. *Psychological Reports, 76*, 1343–1354.

Naglieri, J. A., & Gottling, S. H. (1997). Mathematics instruction and PASS cognitive processes: An intervention study. *Journal of Learning Disabilities, 30*, 513–520.

Naglieri, J. A., & Johnson, D. (2000). Effectiveness of a cognitive strategy intervention to improve math calculation based on the PASS theory. *Journal of Learning Disabilities, 33*, 591–597.

Naglieri, J. A., & Kaufman, A. S. (2008). IDEIA 2004 and specific learning disabilities: What role does intelligence play? In E. Grigorenko (Ed.), *Educating individuals with disabilities: IDEIA 2004 and beyond* (pp. 165–195). New York: Springer Publishing Company.

Naglieri, J. A., & Otero, T. (2011). Cognitive assessment system: Redefining intelligence from a neuropsychological perspective. In A. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 320–333). New York: Springer.

Naglieri, J. A., & Paolitto, A. W. (2005). Ipsative comparisons of WISC-IV index scores. *Applied Neuropsychology, 12*, 208–211.

Naglieri, J. A., & Pickering, E. B. (2003). *Helping children learn: Intervention handouts for use at school and home*. Baltimore: Brookes Publishing.

Naglieri, J. A., & Pickering, E. (2010). *Helping children learn: Intervention handouts for use in school and at home* (2nd ed.). Baltimore: Brookes.

Naglieri, J. A., & Rojahn, J. R. (2004). Validity of the PASS theory and CAS: Correlations with achievement. *Journal of Educational Psychology, 96*,174–181.

Naglieri, J. A., Rojahn, J. R., Matto, H. C., & Aquilino, S. A. (2005). Black white differences in intelligence: A study of the PASS theory and cognitive assessment system. *Journal of Psychoeducational Assessment, 23*,146–160.

Naglieri, J. A., Goldstein, S., Delauder, B. Y., & Schwebach, A. (2006a). WISC-III and CAS: Which correlates higher with achievement for a clinical sample? *School Psychology Quarterly, 21*, 62–76.

Naglieri, J. A., Rojahn, J., & Matto, H. (2007). Hispanic and non-Hispanic children's performance on PASS cognitive processes and achievement. *Intelligence, 35*, 568–579.

Naglieri, J. A., Otero, T., DeLauder, B., & Matto, H. (2007). Bilingual Hispanic children's performance on the English and Spanish versions of the cognitive assessment system. *School Psychology Quarterly, 22*, 432–448.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014). *Cognitive assessment system* (2nd ed.). Austin: Pro-Ed.

O'Donnell, L. (2009). The Wechsler intelligence scale for children. In J. A. Naglieri & S. Goldstein (Eds.), *A*

practitioner's guide to assessment of intelligence and achievement (4th ed., pp. 153–190). New York: Wiley.
Otis, A. S. (1918). *An absolute point scale for the group measurement of intelligence*. https://ia600401.us.archive.org/28/items/cu31924073896650/cu31924073896650.pdf. Reprinted from the Journal of Educational Psychology, IX.
Otis, A. S., & Lennon, R. T. (1979). *Otis-Lennon school ability test*. New York: Psychological Corporation.
Pintner, R. (1923). *Intelligence testing*. New York: Henry Holt.
Pressley, M. P., & Woloshyn, V. (1995). *Cognitive strategy instruction that really improves children's academic performance* (2nd ed.). Cambridge, MA: Brookline.
Prifitera, A., & Saklofske, D. (1998). *WISC-III clinical use and interpretation: Scientist-Practitioner perspectives*. New York: Academic.
Ramirez, R. R., & de la Cruz, G. P. (2002). *The Hispanic population in the United States: March 2002* (Current population reports, P20-545). Washington, DC: U.S. Census Bureau.
Roid, G. (2003). *Stanford-Binet* (5th ed.). Itasca: Riverside.
Scheid, K. (1993). *Helping students become strategic learners*. Cambridge, MA: Brookline.
Schofield, N. J., & Ashman, A. F. (1987). The cognitive processing of gifted, high average, and low average ability students. *British Journal of Educational Psychology, 57*, 9–20.
Silverstein, A. B. (1982). Pattern analysis as simultaneous statistical inference. *Journal of Consulting and Clinical Psychology, 50*, 234–240.
Silverstein, A. B. (1993). Type I, Type II, and other types of errors in pattern analysis. *Psychological Assessment, 5*, 72–74.
Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego: Jerome M. Sattler.
Suzuki, L. A., & Valencia, R. R. (1997). Race-ethnicity and measured intelligence. *American Psychologist, 52*, 1103–1114.
Wechsler, D. (1939). *Wechsler-Bellevue intelligence scale*. New York: Psychological Corporation.
Wechsler, D. (1991). *Wechsler intelligence scale for children* (3rd ed.). San Antonio: The Psychological Corporation.
Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio: The Psychological Corporation.
Wechsler, D. (2005). *Wechsler individual achievement test (WIAT II)* (2nd ed.). London: The Psychological Corp.
Wechsler, D. (2014). *Wechsler intelligence scale for children* (5th ed.). San Antonio: The Psychological Corporation.
Weiss, L. G., Saklofske, D. H., Prifitera, A., & Holdnack, J. A. (2006). *WISC-IV advanced clinical interpretation*. New York: Elsevier.
Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson revised tests of cognitive ability: Standard and supplemental batteries*. Itasca: Riverside Publishing.
Wendling, B. J., Mather, N., & Shrank, F. A. (2009). Woodcock-Johnson III tests of cognitive abilities. In J. A. Naglieri & S. Goldstein (Eds.), *A practitioner's guide to assessment of intelligence and achievement* (pp. 191–232). New York: Wiley.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Itasca: Riverside Publishing Company.
Yerkes, R. M. (1921). *Memoirs of the National Academy of Sciences Volume XV*. Washigton: Government Printing Office.
Yoakum, C. S., & Yerkes, R. M. (1920). *Army mental tests*. New York: Henry Holt and Company.

21 The Relationship Between Theories of Intelligence and Intelligence Tests

W. Joel Schneider and Dawn P. Flanagan

In the field of cognitive ability assessment, the separate spheres of theory, research, and practice are not that separate, evolving in chaotic fits of individuation and rapprochement. Their relationship is complicated: Research can be particularly cruel to theory, theory makes impossible demands of practice, and practice can go for long stretches acting as if research is not even in the room. Yet, they are family and they need each other. Research surprises theory with thoughtful gifts, theory gives sound advice when practice is troubled, and practice helps research to get out of the lab and meet people. Their story is too long and complex to tell in full, but there are helpful ways of making sense of their relationships.

Kamphaus, Winsor, Rowe, and Kim (2012) have presented the history of intelligence test interpretation as a series of waves of new practices. This is, of course, only a metaphor, but it communicates historical truth reasonably well. The metaphor is possibly misleading because it suggests that each wave was discrete and completely displaced the effects of the previous wave. In truth, the "waves" overlapped by many years, and the practices of earlier waves never died out. First-wave practices are alive and well in many applications of intelligence tests, as are second- and third-wave practices. Each wave describes a kind of central tendency of each era, but there are many outliers we could point to as exceptions to the trends. We borrow extensively from the historical analysis and vocabulary of Kamphaus and colleagues, but our parsing of history differs slightly from theirs on several minor points.

Wave 1: General Ability (1904–Present)

After numerous historical anticipations from Galton, J. M. Cattell, and many others, British researcher Charles Spearman (1904) inaugurated the era of scientific models of intelligence. Almost simultaneously, French scholars Alfred Binet and Théodore Simon (1905) published the first well-constructed, reasonably normed, and practical method of identifying children who were in need of special education services because of low intelligence.

Binet (1905) believed that Spearman's theories were overly simplistic, and Spearman (1930) considered Binet's tests to be rather crude. Nevertheless, despite this mutual criticism, Spearman's ideas and Binet's tests aligned well enough to launch a movement. Because of the efforts of IQ evangelists such as Cyril Burt,

W.J. Schneider (✉)
Department of Psychology, Illinois State University, Campus Box 4620, Normal, IL 61790-4620, USA
e-mail: wjschne@ilstu.edu

D.P. Flanagan, Ph.D.
Department of Psychology, St. John's University, Queens, NY, USA
e-mail: flanagad@stjohns.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_21,

Godfrey Thomson, Henry Goddard, and Louis Terman, Binet's tests were translated, refined, and normed in countries all over the world.

For most of the twentieth century, the primary focus of cognitive ability testing was to determine a person's level of general intelligence. Researchers devoted an extraordinary amount of effort exploring diverse theoretical and practical questions about general intelligence and its correlates. This literature is now truly enormous, with hundreds of thousands of studies making use of IQ tests. In contrast with some of the delicate and fickle findings that characterize some subdomains of psychology, the archives of IQ research are replete with robust results that reliably replicate. High IQ is not only predictive of degrees and grades but also wealth and health (Gottfredson 1997).

Because of the early successes of IQ research (and other, less laudable motivations), many institutions began to use them to make decisions about individuals. The use of IQ tests to make high-stakes decisions aroused strong negative sentiments right from the beginning (Chapman 1988):

> At times the rhetoric of the testers was elitist, hereditarian, and defensive. By overselling their reform, they provoked an angry response. By claiming for themselves and their instrument an extraordinary power to predict human destiny, they met sharp resistance from those who believed in America as a place of opportunity, not self-fulfilling prophecies. (p. 176)

Nevertheless, the value of IQ tests to important institutions in education, government, industry, and the military was so great that the use of IQ tests has been resilient to multiple waves of scholarly criticism and public scrutiny (Kaufman 2000).

Intelligence tests were famously involved in helping the U.S. military with an efficient means of finding an initial placement for millions of recruits during World War I. The success of the initial attempts of psychologists to help in the war effort has been disputed (Samelson 1977), but it is clear that the U.S. military still finds value in their continued use. Currently, no one who scores substantially below average on the Armed Forces Qualifying Test (a measure of literacy and numeracy strongly correlated with IQ) is allowed to join the U.S. military. The U.S. military also uses a broader collection of aptitude tests (the Armed Services Vocational Aptitude Battery) to facilitate initial placement of recruits and to qualify applicants for specialized training.

Organizational psychology has a long tradition of using IQ tests for personnel selection. The primary emphasis on general intelligence in personnel selection persists to this day, largely because it has been difficult to demonstrate the differential validity of specific abilities in job performance (Schmidt & Hunter 2004). Commercially available tests of general ability are used by government agencies and large firms to make hiring and promotion decisions. These tests are still used in a manner that is largely consistent with "first-wave" practices.

In educational settings, intelligence tests have been used to identify the extremes of general intelligence so that children can be placed in classrooms deemed to be better suited to their learning speed. In the 1920s, college aptitude tests (close relatives of intelligence tests) began to be used by selective institutions to identify academically talented students who did not have the advantage of having been educated in elite prep schools.

The focus on the general factor of intelligence is alive and well both in clinical research and in practice. Many scholars believe that the clinical use of intelligence tests should focus primarily on the general factor of intelligence as measured by overall IQ. These scholars do not deny that other factors exist or that they are associated with important outcomes (Glutting et al. 2006). However, they argue that we measure few abilities with sufficient validity to use such measurements to make helpful decisions about individuals (Canivez 2013). These conclusions are, of course, passionately disputed by many scholars and practitioners in the field. Nevertheless, despite the vituperations of partisans, the research is still ambiguous enough that either position is still intellectually respectable (Schneider 2013a).

Wave 2: Clinical Profile Analysis (1940s–Present)

Almost as soon as Spearman proposed his theory of general intelligence, researchers (including Spearman himself) began to produce evidence that broad abilities other than *g* exist (Spearman 1927). Although the research demonstrating the existence of multiple broad factors was extremely sophisticated, the application of this knowledge in practice was somewhat haphazard and clumsy at first. Practicing psychologists began to use multiple scores from IQ tests to better understand their evaluees, but they tended to use insights drawn from clinical lore rather than from demonstrated research findings.

Practitioners were trained to evaluate ability profiles in what today is known to be a psychometrically naïve manner, without fully accounting for score reliability and stability. Unrealistically specific test profiles were said to be characteristic of specific psychiatric diagnoses. It would be wonderful if IQ tests were capable of making such useful distinctions. Unfortunately, although it is true that certain diagnostic groups do have weak tendencies toward certain kinds of profiles, the positive predictive power of such profiles to identify specific disorders is very low (Watkins et al. 1997).

We should not be too quick to cast stones at second-wave scholars and their experientially and rationally derived clinical lore. Though it is true we do not live in houses made entirely of glass, our homes still have many large and fragile windows. Furthermore, those windows are not merely decorative but provide useful views of reality. Much of what practitioners still do is necessarily based on rationally derived principles, clinical experience, common sense, and basic human dignity. It would be desirable to reinforce those glass panes with strong research findings, but true progress is difficult and thus slow.

Much of what is called "naïve psychometrics" can be considered empathic wisdom on the part of practitioners. Among the most thoughtful and articulate scholars to promote second-wave practices were Rapaport, Gil, and Schafer (1945–1946). Their approach was to first look at an evaluee's scores but also to look between the scores, beneath the scores, and beyond the scores to assess psychological processes operating idiosyncratically in the evaluee:

> First, we found that one must consider not only every subtest score, but every single response and every part of every response, as significant and representative of the subject. Naturally, many of the intelligence-test responses are highly conventionalized and that a subject knows who was president of the United States at a particular time merely adds to his general score. But where the response deviates from the conventional, the deviation does not merely fail to add to his score; it must also be considered as a characteristic which may give us material toward the understanding of the subject….
>
> Second, we found that one may gain some understanding of the subject by comparing the successes and failures on a given type of test item. Thus, if a subject knows how many pints there are in a quart, but does not know what the Koran is, this will give us merely an idea of his range of information. But if he knows what the Koran is and asserts that a quart has four pints, we must consider the presence of a temporary inefficiency, and if he insists that the capital of Italy is Constantinople or that the Vatican is a robe, psychotic maladjustment will have to be considered….
>
> In general, one might say that this approach to intelligence testing requires a very different attitude toward tests than does routine intelligence testing, which hinges upon correct appraisal of whether a response is to be considered passing or failing…. On the part of the tester, it requires a great deal of attention to any type of deviation from the usual run of test performances. (pp. 67–68)

The second-wave tradition was extended by many scholars, including Kaufman (1979), Sattler (1974), and especially Kaplan with her Boston Process Approach (Milberg et al. 1996). Recent advances in this tradition can be found in many sources (e.g., Hale and Fiorello 2004; McCloskey et al. 2008).

Wave 3: Psychometric Profile Analysis (1970s–Present)

Although research on cognitive ability had always made use of sophisticated statistical methods, practitioners were rarely trained to take advantage of them. In the years following World War II, practical textbooks on psychometrics began to be published (e.g., Gulliksen 1950), factor analytic studies of the Wechsler scales showed that non-*g* abilities could be measured (Cohen 1959), and findings from neuropsychology showed that estimating these non-*g* abilities was not a waste of time (Halstead 1947).

The early writings of Sattler (1974) and Kaufman (1979) were extremely persuasive in showing that it was worthwhile to use psychometrics to extract meaningful information from test scores beyond the *g* factor. Not only were Sattler and Kaufman persuasive and practical, their comprehensive and “intelligent” approach to intelligence test interpretation made the process interesting and meaningful to practitioners. Sattler’s and Kaufman’s early writings were not purely psychometric nor were they atheoretical, but compared to later efforts (including their own), they were less theoretically driven. Much of their earlier work was devoted to retrofitting existing clinical measures not designed to cleanly measure any theoretical entities other than *g* with theoretical interpretations derived from Guilford’s (1967) structure of-intellect model and Osgood’s (1963) psycholinguistic model. Even at the time, these connections to such theories were made with ample warning of their speculative nature.

Wave 4: Application of Theory to Interpretation (1980s–Present)

In third-wave approaches, clinical measures were mapped onto then-current theories of intelligence; in fourth-wave approaches, new clinical measures were created (or old ones adapted) to operationalize successful theories of intelligence. Such measures had existed for many decades, but they were almost exclusively used for research (e.g., Thurstone’s Primary Ability Battery).

The writings of Soviet neuropsychologist Alexander Luria (1966, 1973) excited a new generation of cognitive ability researchers and test developers. Because of the tremendous number of casualties suffered by the Soviets during World War II, Luria was able to study and treat a large number of men with diverse, and in many cases focal, brain injuries. The discipline of neuropsychology made rapid advances during this time.

Luria had legendary clinical skills, often making up new neuropsychological tests on the spot to test hypotheses about individual evaluees. Although the creative genius of Luria’s approach cannot be duplicated, many of his ideas were later used by others. The Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman 1983), the Cognitive Assessment System (CAS; Naglieri & Das 1997), and the Luria-Nebraska Neuropsychological Battery operationalized portions of Luria’s methods and theories. Das, Naglieri, and Kirby (1994), in particular, developed and extended Luria’s theories with their planning, attention, simultaneous, and successive (PASS) theory and operationalized PASS constructs with the CAS.

Although Cattell’s Gf-Gc theory had been proposed in 1941, it was not tested directly until the 1960s with the help of Cattell’s student, John Horn (Cattell 1963; Horn & Cattell 1966). Horn and Cattell, sometimes separately, sometimes together, refined the theory and subjected it to critical tests throughout the 1960s and 1970s. By the late 1980s, the evidence for the theory was so strong that Woodcock restructured his battery of tests to be an admirably comprehensive operationalization of Gf-Gc theory (Woodcock & Johnson 1989). The Stanford-Binet Intelligence Scale: Fourth Edition (SB:FE; Thorndike et al. 1986) and the Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman 1993) were also partial operationalizations of Gf-Gc theory. Gf-Gc theory’s status was greatly enhanced when Carroll’s (1993) massive reanalysis of the world’s cognitive ability research demonstrated persuasively that the model of broad abilities specified by the Horn-Cattell model was largely correct. Nearly all subsequently published multifactor cognitive ability tests have been based explicitly or implicitly on the ideas of Cattell, Horn, and Carroll (Alfonso et al. 2005), with Naglieri and Das’s (1997) CAS as the primary exception. A number of scholars see considerable merit in both CHC theory and

PASS theory and believe that a successful integration of the two would be of great theoretical and practical benefit (Flanagan, Alfonso, & Dixon 2014; Kaufman & Kaufman 2004).

With these theoretical updates to test batteries, new frameworks for test interpretation were articulated. Building on the work of Woodcock (1990), McGrew (1997) proposed a unified framework for the separate but closely related theories of Cattell, Horn, and Carroll, later termed *CHC theory*. Also building on Woodcock's (1990) pioneering work, the *cross-battery approach* (Flanagan & McGrew 1997; Flanagan et al. 2013b) was developed to help practitioners combine and interpret information from multiple test batteries in a psychometrically defensible manner according to CHC theory.

What characterizes fourth-wave test interpretations is the confidence with which tests can be interpreted according to well-developed theories such as CHC theory and PASS theory. The confidence is reasonably warranted because of many replicated findings in cross-battery confirmatory factor analyses (Flanagan et al. 2013a; Reynolds et al. 2013). Instead of being bound to one battery or another, practitioners can flexibly pick tests from any battery to meet specific situational demands.

Psychometric Forerunners of CHC Theory

As shown in Fig. 21.1, CHC theory has three parents (g_f-g_c theory, extended Gf-Gc theory, and three-stratum theory) and at least two grandparents (two-factor theory and primary mental abilities). In addition, it has two important first cousins once removed (hierarchical Group Factor theory and triadic theory).

Two-Factor Theory of Ability

Spearman's (1927) two-factor theory of ability is the simplest theoretical justification for creating a composite score from diverse question types. The two-factor theory assumes that, aside from error, test scores are determined by two kinds of ability: general and specific. General ability

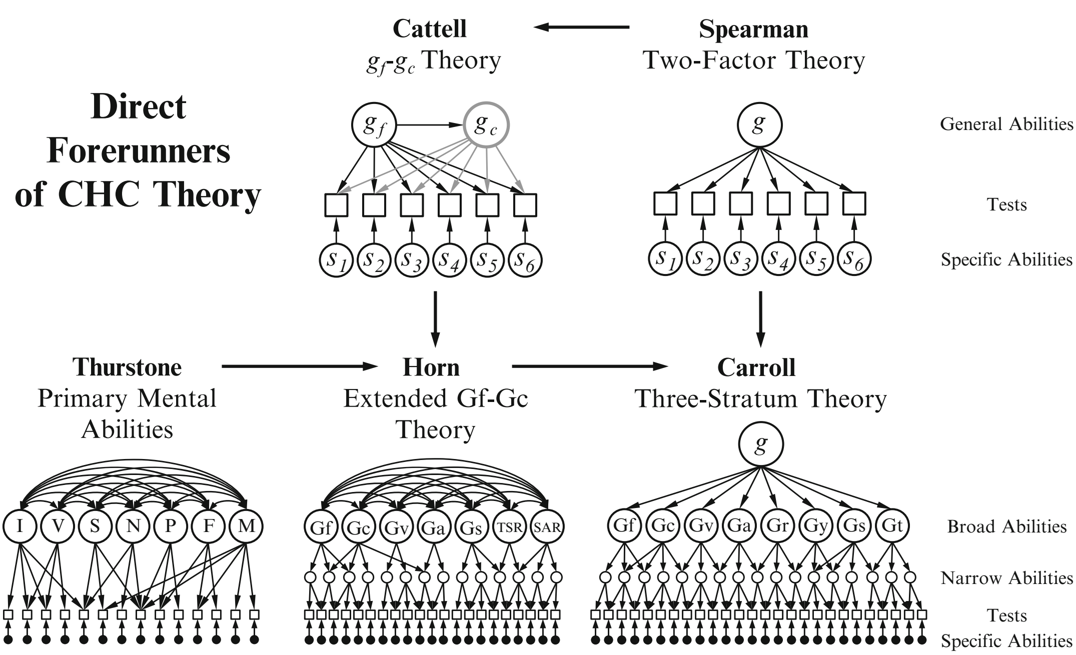


Fig. 21.1 Direct forerunners of CHC theory

(g) influences all test scores, whereas specific abilities influence just one kind of test each, as shown in Fig. 21.2. The reason that Spearman believed that the two-factor theory had merit is that in many data sets, as predicted by the theory, the product of the g loadings of any two tests was close to the observed correlations. For example, if two tests have g loadings of 0.7 and 0.8, respectively, then the two-factor theory predicts that their observed correlation will be 0.56.

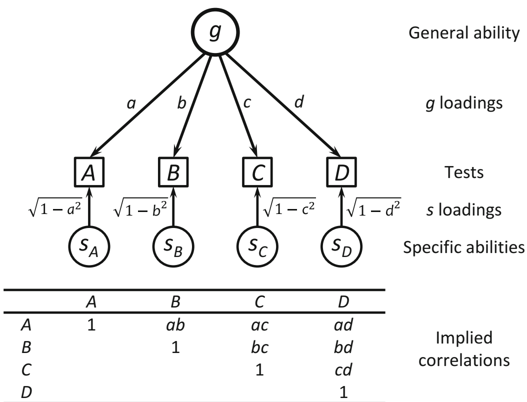


Fig. 21.2 Spearman's two-factor theory of ability. *Note*: All observed and latent variables are depicted as z-scores

One of Spearman's many insights was that even if g has a small influence on any particular test item, g can be estimated reasonably well by adding many test scores together. In Fig. 21.3, it can be seen that the sum of two ability tests is a better measure of g than either of the two tests would be if used by themselves. This provides test developers with a rationale for making IQ tests that select from a wide variety of test formats, contents, and task demands.

Spearman's greatest contribution to CHC theory, surprisingly, is not the theoretical entity g. Galton's (1888) invention of the correlation coefficient was a major breakthrough, allowing researchers to ask questions about pairs of variables instead of just one variable at a time. What Galton did with pairs of variables, Spearman did with whole sets of variables, simultaneously. Spearman moved beyond bivariate correlations to the evaluation of whole correlation matrices. That is, Spearman directed our focus away from piecemeal analyses to omnibus evaluations of a theoretical model as a whole. Spearman's inven-

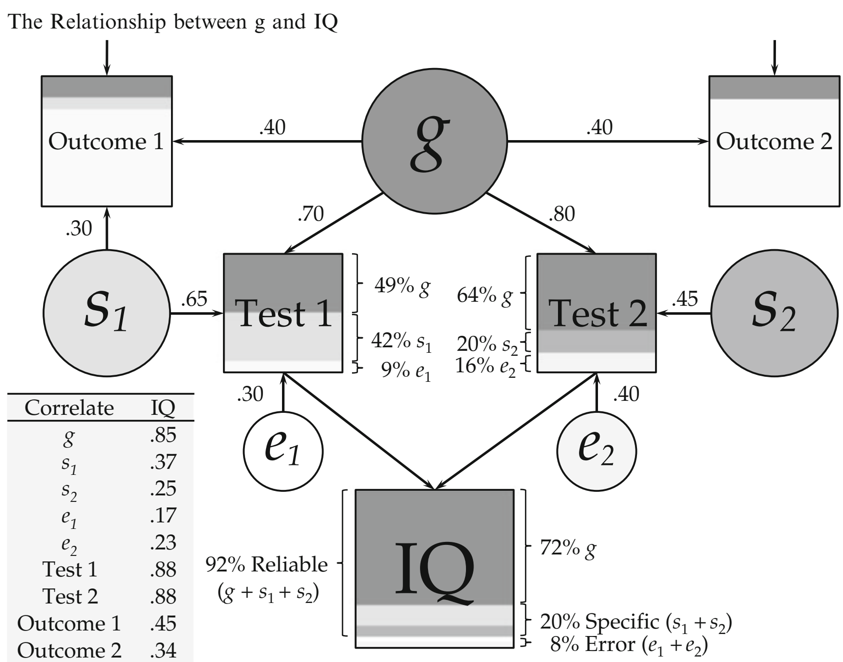


Fig. 21.3 The relationship between g and IQ

Table 21.1 Similarities across psychometric models of ability

Primary abilities Thurstone	Triadic theory Cattell	Extended Gf-Gc Horn	Three-stratum Carroll	Cattell-Horn-Carroll
Verbal	Crystallized intelligence	Acculturation knowledge	Crystallized intelligence	Comprehension knowledge
Reasoning	Fluid intelligence	Fluid intelligence	Fluid intelligence	Fluid reasoning
Perceptual speed	General perceptual speed	Cognitive speed	Broad cognitive speediness	Processing speed
		Correct decision speed	Reaction time decision speed	Reaction and decision speed
Word fluency	General retrieval capacity		Broad retrieval ability	Retrieval fluency
Memory	General memory capacity	Tertiary storage and retrieval	General memory and learning	Learning efficiency
	Short-term apprehension and retrieval			Short-term memory
Space	Visualization capacity	Visualization and spatial orientation	Broad visual perception	Visual processing
	Auditory capacity	Listening and hearing	Broad auditory perception	Auditory processing
Number				Quantitative knowledge

tion of a procedure that could do this (a rudimentary version of factor analysis) is impressive enough, but the real breakthrough was the idea that one could and should do this. In a sense, all individual difference researchers are still playing the game that Spearman invented.

Primary Mental Abilities

Spearman's rudimentary factor analysis was a powerful tool, but it required researchers to extract only one factor from a correlation matrix at a time, with each step requiring hundreds of calculations. Spearman's methods could account for group factors, but the calculations were tedious. Thurstone (1934) invented multiple factor analysis, which allowed all common factors to be extracted simultaneously via matrix algebra, facilitating researchers' ability to test more complex models of intelligence. Thurstone (1938) did not deny the existence of *g* but emphasized a small number of what he called *primary mental abilities*, which are neither general across all tests nor specific to each test (see Fig. 21.1).

Thurstone (1938, 1947) conducted an extraordinary amount of research developing a comprehensive battery of over 50 cognitive ability tests which he gave to diverse samples at different ages. Many of his tests are still used in research, but no clinical instruments were explicitly based on Thurstone's primary mental abilities battery. Thurstone's list of abilities varied somewhat from study to study, but a stable subgroup of ability factors was found in almost every study. As seen in Table 21.1, these factors are in most cases directly analogous to broad abilities in Cattell's (1987) triadic theory, Horn's extended Gf-Gc theory (Horn & Blankson 2005); Carroll's (1993) three-stratum theory, and Cattell-Horn-Carroll theory (McGrew 2009; Schneider & McGrew 2012). Thurstone's research identified a number of narrower primary abilities not shown in Table 21.1 such as vocabulary, reading comprehension, induction, deduction, and closure (Guilford 1972), all of which are represented in CHC theory. Thus, CHC theory is not so much a new theory but an elaboration of very robust findings that were first discovered by Spearman, Thurstone, and many other early researchers.

g_f-g_c Theory

Even in his first paper on general intelligence, Spearman (1904) was open to the possibility that there were multiple general factors. Spearman's student, Raymond Cattell, found evidence for two general factors of intelligence (Cattell 1941, 1943) which he called fluid intelligence (g_f) and crystallized intelligence (g_c). What athletic talent is to the body, fluid intelligence is to the brain. It represents the speed, power, efficiency, and overall integrity of the cerebral cortex. In Cattell's (1987) thinking, g_f is not an ability itself but an influence on many abilities, particularly those abilities that require controlled attention and on-the-spot problem solving. g_c is acquired knowledge, particularly information stored in declarative memory. Whereas g_f is easily damaged (or even temporarily disrupted by fatigue or intoxication), g_c is relatively robust to the effects of disease, aging, and injury.

Extended Gf-Gc Theory

Under Cattell's (1963) supervision, John Horn conducted the first empirical test of Cattell's theory of intelligence. Some predictions held up well, but the theory needed to be refined and expanded. The new extended Gf-Gc theory (Horn and Cattell 1966) did not just borrow Thurstone's primary abilities, it also provided Thurstone's ever-lengthening list with some much needed order. That is, it showed that Thurstone's primary abilities had a higher-order structure and that different primary abilities were differentially subject to Gf, Gc, and other general factors such as Gv (general visualization), Gs (general speediness), and Gr (general memory fluency). Later, broad abilities were added by Horn's research independent of Cattell. Cattell's own theorizing and research also continued independently of Horn, though the two continued to publish together from time to time. Confusingly, the term *g_f-g_c theory* refers to Cattell's (1941; 1943) original idea, and the *(extended) Gf-Gc theory* (aka the *Horn-Cattell model*) was refined mostly by Horn (1985). As described below, Cattell's later work is known as the triadic theory of ability.

Three-Stratum Theory of Cognitive Abilities

Carroll (1993) systematically reanalyzed every data set he could find that would help him generate a new taxonomy of abilities. After evaluating the results, he found that he had largely reconstructed Horn's model with a number of important differences. Carroll believed that there was strong evidence for the *g* factor, whereas Horn did not. As seen in Table 21.1, Horn and Carroll parse memory differently. Neither theorist is wrong; they simply have different emphases. It is true that nature must be carved at its joints, but different carvers can find equally valid ways to slice complex domains.

Triadic Theory of Ability

Cattell's (1971, 1987) later work is called the *triadic theory of ability* (see Fig. 21.4). He hypothesized that some mental capacities (fluid intelligence, memory, retrieval fluency, and processing speed) are not localized in specific regions of the brain. These *general capacities* depend on the overall integrity of the cortex.

Some mental capacities, called *provincial powers*, mostly depend on regions of the brain associated with specific sensory modalities. For example, individual differences in visualization capacity result from better functioning in the primary and secondary visual association cortex regions in the occipital lobe and the rear portions of the parietal and temporal lobes. Each sensory modality with a specific region of primary and secondary association cortex is hypothesized to have its own provincial power. At the time, Cattell had good evidence for the existence of visualization and auditory capacity. He hypothesized that other sensation-specific ability factors would be isolated by future research (e.g., olfactory, gustatory, tactile, and cerebellar).

Whereas general capacities and provincial powers represent influences on abilities, *agencies* are developed abilities. In the triadic theory, capacities like fluid intelligence are not true abilities but are features of a person's brain that put

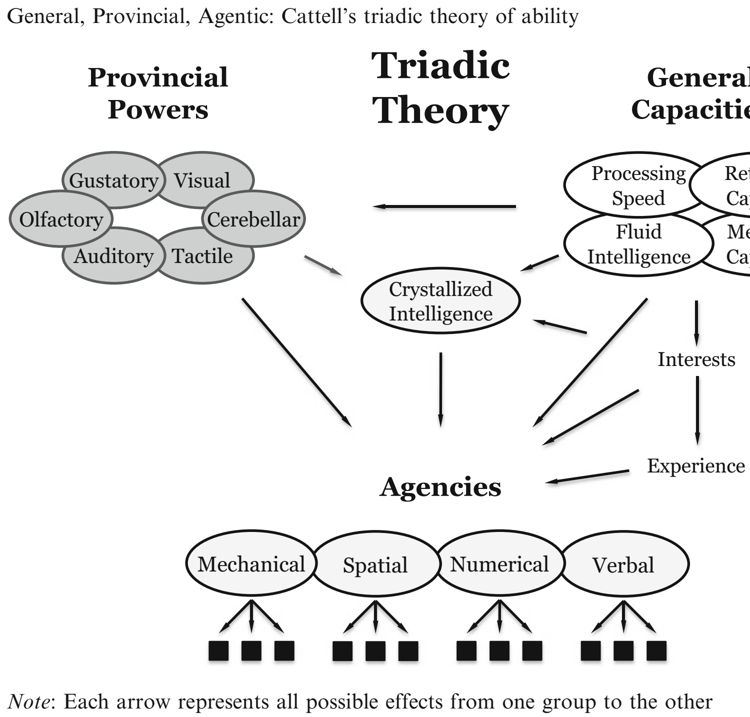


Fig. 21.4 General, provincial, agentic: Cattell's Triadic Theory of Abilities. *Note*: Each *arrow* represents all possible effects from one group to the other

limits on the sorts of developed abilities a person can acquire. The influence of general capacities and provincial powers is always indirect and thus can only be inferred. Therefore, all abilities observed directly are learned abilities. There are many agencies (more than what is shown in Fig. 21.4), but one of them (crystallized intelligence) has such a large influence on so many tasks that it acts as if it were a general capacity.

Because of the success of g_f-g_c theory (Cattell 1943; Horn and Cattell 1966), it is not a hard sell to suggest that Cattell should be read in the original. The rewards of doing so are unexpectedly fabulous. Cattell was a remarkable communicator and a dazzlingly creative scholar. Amazingly, Cattell's primary research focus was personality, not intelligence. In praise of Cattell's many-splendored talents, Carroll (1984, p. 300) noted that Cattell's work on abilities was "a sideline, conducted, as it were, with his left hand." Consider this review by Jensen (1974) of Cattell's (1971) *Abilities: Their structure, growth, and action*:

> Looking back over the history of the field, I would judge this book to be the single most ambitious and original contribution to the study of mental abilities since Charles Spearman's *The Ability of Man* (1927). Cattell's book is simply unrivaled in its scope. Whatever the reader's overall assessment of the book's virtues and faults, few who study it thoroughly will fail to think of it as the work of a brilliant and encyclopedic intellect. (p. 291)

It feels odd to wholeheartedly agree with Jensen about anything, but we are happy to make an exception in this case. We also agree with Jensen's evaluation of the larger 1987 sequel, *Intelligence: It's Structure, Growth, and Action*, written when Cattell was a spry 82 years of age. It, too, is a remarkable summary of Cattell's thinking but is only barely updated from 1971; many references from the 1950s and 1960s are referred to as "recent." Furthermore, Cattell seems to have been so busy with his own research that he does not seem particularly aware of anyone else's contributions to the field after the 1960s. Nevertheless, Cattell's 1971 masterwork on intelligence and its 1987 sequel are filled with page after page of insights and fascinating details that are rarely mentioned anywhere else by anyone else. For example, his explanation of how crystallized intelligence becomes functionally autonomous from fluid intelligence is detailed and remarkably ingenious (Cattell 1987, pp. 146–152).

Hierarchical Group Factors

The work of British psychologist Philip Vernon never disappeared from view, but it is now resurgent because of the recent successes of its successor, the verbal-perceptual-image rotation model (VPR; Johnson and Bouchard 2005). Most descriptions of Vernon's model point out that between the *g* factor and more specific abilities, abilities can be broadly categorized as verbal/educational (v:ed) or spatial/mechanical (k:m). Although not an explicit feature of Carroll's model, this was a distinction Carroll (1993, p. 60) believed was valid, noting that there are probably many such intermediate strata in the structure of ability. The later elaborations of Gf-Gc theory by Horn and Blankson (2005) and by Cattell (1971) make similar (though not identical) distinctions.

A common misperception of Vernon's model is that all abilities are subsumed by the v:ed/k:m clusters. Not at all! Vernon (1961) believed that a separate auditory and musical talent cluster of abilities (similar to Ga) was distinct from both v:ed and k:m (see Fig. 21.5). He also believed that sensory discrimination factors and various forms of imagery were also distinct. This is but one way in which Vernon's model has been lamentably oversimplified.

At first glance, Vernon's model looks a lot like many other hierarchical models. However, there is an important difference. Vernon's model is not a straight reporting of factor analytic results. Like Carroll's model, it is instead a summary judgment by Vernon of many factor analyses. In most hierarchical models of ability, including Carroll's, the hierarchies are arranged neatly into two or three strata with each lower-order ability having one and only one "parent" ability. In Vernon's model, the hierarchy is not neatly arranged, having a vaguely organic appearance. Abilities diverge and reconverge at different levels of the hierarchy, with some abilities having multiple "parents."

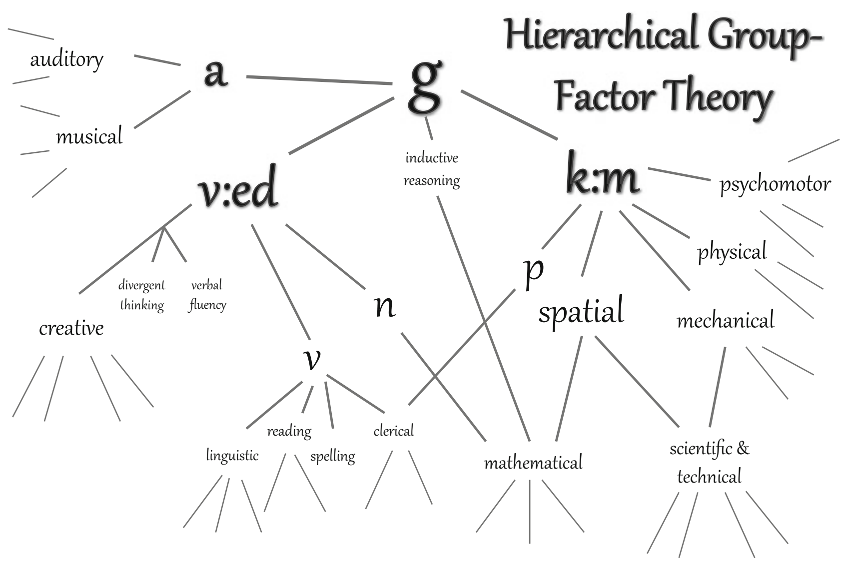


Fig. 21.5 Vernon's hierarchical group factor theory

This is incredibly sensible and important, particularly with respect to the achievement factors, which are clearly multiply determined.

As Horn and Noll (1997, p. 84) famously warned, a mature theory of cognitive abilities is unlikely to have the rectilinear system of factors implied by factor analysis, but is likely to be described by the "rounded and irregular structures of mother nature." As an example of an overly regular model, early diagrams of CHC theory depict reading/writing ability (Grw) and quantitative ability (Gq) as being connected to other factors solely by *g*, though this was never what verbal descriptions of the theory implied (McGrew 1997, 2005, 2009). Grw and Gq were always hypothesized to have direct relationships with Gc, Gf, and other abilities. Nevertheless, this confusion led Major, Johnson, and Deary (2012) to test and reject the models implied by the misleading figures.

Before Major, Johnson, and Deary's paper was published, steps had already been taken to provide a clearer picture of CHC theory, showing that some broad abilities were more closely related than others (Schneider & McGrew 2012). Further clarifications of the exact nature of the relations among CHC broad abilities are still needed. As it stands, CHC theory and the VPR model are now explicitly more alike in several ways (though not identical). If the VPR model continues its run of empirical successes, practitioners will have increasing reason to explore its forerunner, Vernon's model.

Cattell-Horn-Carroll Theory of Cognitive Abilities

If we have the separate works of Cattell, Horn, and Carroll, why do we need the Cattell-Horn-Carroll theory of cognitive abilities? First, Horn and Carroll produced very similar theories in part because they were based on the same data sets. They differ primarily in that Horn's model has no place for *g*, whereas Carroll puts *g* in a place of prominence in the hierarchy. Woodcock and McGrew recognized that the debate about the validity of the *g* factor was not likely to end soon. They wanted a label for the common ground on which both models stood and both Horn and Carroll agreed to have their models yoked together under the CHC banner (McGrew 2005). In CHC theory, *g* is present but with a de-emphasized and uncertain status. This has worked to create an engaged community of scholars, researchers, and practitioners who can talk about cases, discuss research findings, and suggest refinements to the model without having to refight constant battles about the existence of *g*.

Second, now that all three original theorists have passed away, their models will calcify if not continuously updated. Without the CHC label and the ecumenical space it creates, each scholar with a slightly different opinion (i.e., everyone) might feel the need to rename the model, causing unnecessary confusion and fragmentation (e.g., Schneider's revised McGrew-Horn-Flanagan model). Since CHC theory was introduced (McGrew 1997), it has been updated twice in response to new research (McGrew 2009; Schneider & McGrew 2012). It is likely that more updates will follow as needed.

CHC theory recognizes *g* at the top of the hierarchy, 16 broad abilities below *g*, and about 80 narrow abilities nested within the broad abilities. In several places, there are intermediate categories between the three strata. Some parts of the taxonomy are more settled than others. CHC theory has largely maintained Carroll's (1993) nomenclature, with departures from the original detailed in Schneider and McGrew (2012).

In Fig. 21.6, there are "level" factors (abilities defined by the difficulty of the task) in the top row and "speed" factors (abilities defined by the rate at which simple tasks can be completed). However, the distinction between level and speed is not a true dichotomy because there are speed-accuracy trade-offs in most tasks. It is not known whether the speed factors are merely descriptively similar or if there is a superordinate speed factor that is distinct from *g*. Some aspects of cognitive speed are difficult to measure distinctly; for example, some rapid naming tasks likely involve a murky mix of perception speed, memory retrieval fluency, oral-motor articulation speed, and attentional fluency.

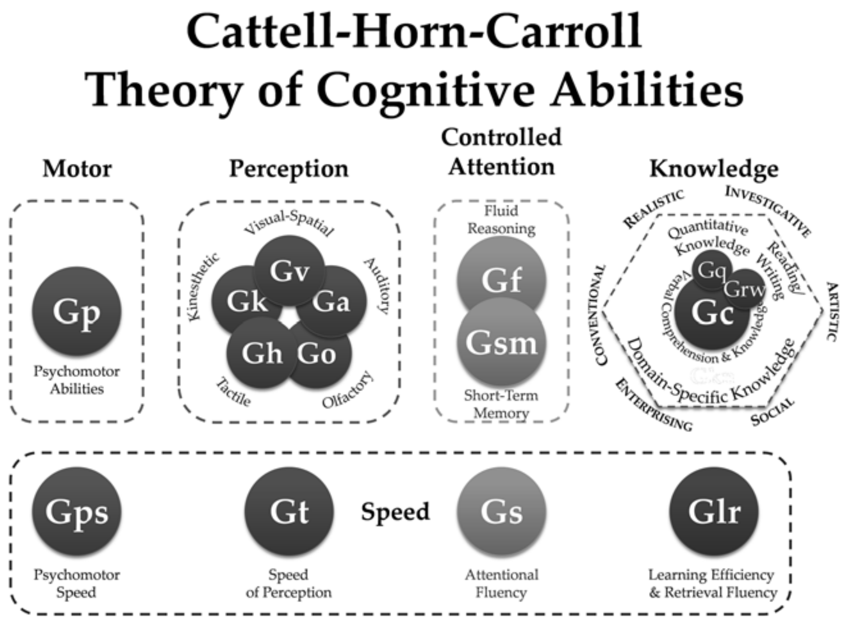


Fig. 21.6 Broad ability groups in CHC theory

The perceptual abilities in Fig. 21.6 are grouped together for conceptual economy rather than for structural reasons. That is, they do not correlate more with each other than with abilities in other categories. Each broad perceptual ability refers to the efficiency and power of perceptual reasoning a person is capable of within a particular modality. It is not known whether each perceptual ability will require its own speed factor or if there are common constraints on perceptual speed.

The connections among the knowledge abilities in Fig. 21.6 are not merely descriptive. For example, language ability and general knowledge directly influence reading comprehension and written expression. Gc and Gkn (domain-specific knowledge) are not sharply distinct abilities, and Gkn is not "an ability" at all but a diverse, fragmented collection of abilities. Gc and Gkn represent two ways in which knowledge serves different functions. Gc is broad knowledge that facilitates communication across an entire culture. Gkn is specialized and deep knowledge that is primarily useful within a particular field, often an extremely narrow one. It is unlikely that the brain stores information about Charlemagne, for example, in different places or in fundamentally different ways depending on whether one is a historian. However, the densely interconnected networks of association in expert knowledge make possible a phenomenon called *expertise wide-span working memory* (Horn and Blankson 2005), in which an expert can hold in mind and process much more information in working memory than is otherwise possible with information outside the expert's specialty. Gkn is hexagonally shaped in Fig. 21.6 because specialized knowledge is partly influenced by career interests, which do have this structure (Holland 1985). However, the hypothesis that domain-specific knowledge has the same structure as career interests is still untested, and the analogy between the two domains is likely to be only approximate.

In early work on fluid intelligence, what are now termed *working memory capacity* tests (tests that require simultaneous short-term storage and processing of information) were used as markers of fluid ability. Although this practice has been discontinued, the connection between Gsm and Gf is intimate (Kane et al. 2004), most likely because of a shared dependence on the ability to control the focus of attention (Unsworth &

McMillan 2014). At the core of fluid intelligence is the ability to perceive new relationships between objects and concepts held temporarily in working memory. Better working memory capacity is not synonymous with better reasoning capacity, but reasoning is much facilitated by working memory capacity.

Current conceptions of working memory (Baddeley 2006; Unsworth & Engle 2007) emphasize the tight connection between short- and long-term memory. In Fig. 21.6, Gsm and Glr are misleadingly far apart. In both Cattell's triadic model and Carroll's model, they are grouped in the same superordinate category, *memory*. Clean measures of any one memory process are probably impossible to design. All memory tasks involve attention, encoding, processing, storage, and retrieval but are designed with different information processing bottlenecks so that different mixtures of memory abilities determine performance.

Doing Wave 1 Well

On many occasions, Spearman (1915) took Binet and other test developers to task for creating IQ tests with diverse item types but denying the existence of *g*. If all test scores are determined solely by specific abilities, it is hard to justify the practice of calculating IQ. Why? If only specific abilities are measured, then a person's IQ would depend on the arbitrary choices made by test developers as to which abilities are measured. Different IQ tests measuring nonoverlapping sets of specific abilities would not correlate with each other.

However, Spearman's objections can be overcome if we posit the existence of a relatively small number of abilities that are broad (i.e., neither general nor specific) and predict a wide variety of criteria, such as academic and occupational success. For example, Cattell's 16PF had a single-score intelligence measure even though Cattell did not believe that there was a single entity called *g*. This practice is justifiable if IQ is seen not as a measure of Spearman's (1904) notion of *general intelligence* but as Binet's notion (Tuddenham 1962) of *intelligence in general*. If there are a few broad abilities that influence performance in many tasks (e.g., verbal ability, spatial ability, logical reasoning, and so forth), a simple average of these important abilities is a useful summary of a person's capacity to act intelligently. This is true whether *g* exists or not.

IQ Versus Weighted Predictors

A finely tuned multiple regression equation with multiple abilities as predictors typically forecasts outcomes somewhat better than IQ (an equally weighted composite score). However, the advantage of a specific multiple regression equation over IQ is not usually large, typically in the range of 1–6 % additional variance explained (Carretta & Ree 2000; Hunter 1986). This fact has long been known and, under plausible conditions in the ability domain is mathematically inevitable (Wainer 1976; Wilks 1938). Dramatically better predictions will come from identifying dramatically better predictors; fine-tuning regression coefficients with existing predictors will only improve predictions a little bit.

What little advantage finely tuned regression equations might have had over equally weighted IQ scores has been reduced because abilities that have the highest correlations with important outcomes tend to be overrepresented in IQ batteries. For example, there are ten subtests in the WISC-IV FSIQ. From the perspective of CHC theory (McGrew 1997, 2005, 2009; Schneider and McGrew 2012), there are three tests of verbal comprehension and knowledge (information, vocabulary, and similarities), two tests of fluid intelligence (matrix reasoning and picture concepts), two tests of working memory (digit span and letter-number sequencing), two tests of processing speed (coding and symbol search), and one test of visual-spatial processing (block design). This unequal mix of abilities (3 Gc:2 Gf:2 Gsm:2 Gs:1 Gv) approximates the relative sizes of the regression coefficients one would find in a multiple regression equation prediction of academic and occupational success.

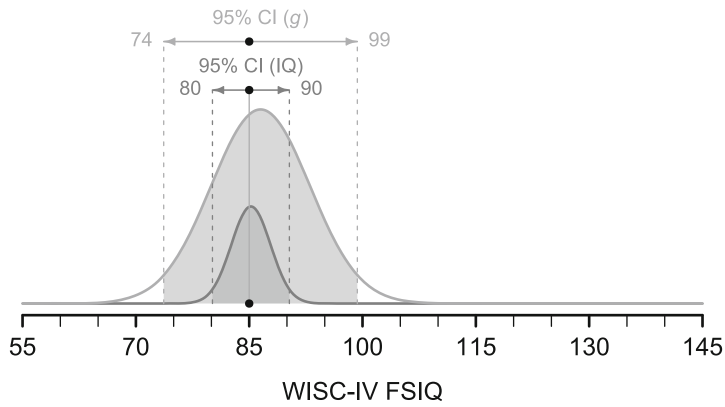


Fig. 21.7 Confidence intervals for *g* and FSIQ when WISC-IV FSIQ = 85

Finally, there is a practical advantage of IQ scores: It is an inconvenient truth that a different multiple regression equation must be computed for each outcome. An IQ score, which is only slightly suboptimal as a predictor, must be computed only once. Thus, whether one believes in *g* or not, an IQ score is still a remarkably effective and practical predictor of important outcomes.

The Relationship Between IQ and *g*

Although the terms "IQ" and "*g*" are sometimes used interchangeably, they are not quite the same. At first glance, their relationship is quite simple: Observed IQ is an imperfect estimate of the latent variable *g*. In truth, IQ can be a complex mix of many things besides *g*. For example, in Fig. 21.3, the IQ score is influenced by the specific abilities measured by Tests 1 and 2. If those specific abilities have correlations with an outcome, the IQ score will have a greater correlation with the outcome than would a more pure estimate of *g*. For example, in Fig. 21.3, both Outcome 1 and Outcome 2 are equally correlated with *g* ($r = .40$). However, because Outcome 1 is also correlated with the specific ability in Test 1, IQ is more strongly correlated with Outcome 1 ($r = .45$) than with Outcome 2 ($r = .34$). Unfortunately, unless the correct analysis is done, *g* will likely get all the credit in its prediction of Outcome 1.

IQ scores are famously reliable—and reliability, famously, is not the same as validity. High reliability means that a *true score* can be located with precision. However, an IQ's true score is not so much an estimate of *g* as it is an estimate of the entire mix of reliable sources of variance. The WISC-IV FSIQ, across age groups, has a classical test reliability coefficient of 0.97. Using a model of the WISC-IV developed by Weiss, Keith, Zhu, and Chen (2013), it can be deduced that *g* makes up 81 % of the variance in the FSIQ, with 16 % of the variance attributed to other abilities and 3 % of the variance attributed to error (Schneider 2013b).

Does this distinction between *g* and the true score matter? It does if you think that FSIQ is an estimate of *g*. Consider a person with a WISC-IV FSIQ of 85. In Fig. 21.7, it can be seen that the 95 % confidence interval for the true score is between 80 and 90, a 10-point range. In contrast, *g* cannot be estimated with the same level of precision; its 95 % confidence interval is between 74 and 99. This is 2.5 times larger than the confidence interval for the true score! Although a 25-point confidence interval covers a wide range, it is not so wide that the estimate is useless; in this case, the estimate allows us to be confident that the person's *g* is lower than the population mean but not extremely so.

Not All General Ability Scores Estimate *g* Equally Well

Although most cognitive batteries estimate *g* reasonably well, some do it more precisely than others. The Stanford-Binet 5 is especially noteworthy in that it has many highly *g*-loaded subtests. Based on a simple 1-factor model using the standardization sample norms (ages 11–16), it is estimated that *g* explains 92 % of the variance in FSIQ (more realistic models result in slightly lower estimates around 90 %). When the SB5 FSIQ is used to estimate *g*, the 95 % confidence interval is only 16 points wide, noticeably more precise than the 25-point interval on the WISC-IV.

It can be estimated that about 72 % of the variance in the WJ III GIA Extended is due to *g*, resulting in a 95 % confidence interval for *g* that is 31 points wide. The reason that the SB5 is so good at estimating *g* and that the WJ III is not is that the SB5 fails to measure any other factor with precision (Canivez 2008). In contrast, the WJ III was explicitly designed to measure abilities with low *g* loadings. Note, however, the fact that the WJ III GIA Extended is less correlated with *g* does not mean that it is less reliable or a worse predictor of important outcomes. On the contrary, it is about as reliable as the SB5 FSIQ, and it predicts outcomes with non-*g* influences. Thus, outcomes that are influenced by *g* alone (e.g., Outcome 2 in Fig. 21.3) are best predicted by highly *g*-loaded tests like the SB5. The WJ III shines in its ability to predict outcomes explained by abilities other than *g* (e.g., Outcome 1 in Fig. 21.3).

Taking Wave 2 Practices Seriously

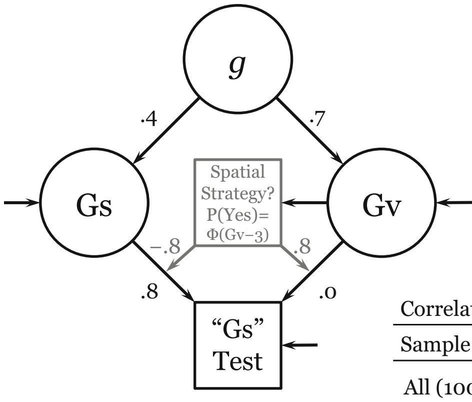
There will never be a complete list of the manifold influences on test performance nor should there be. If someone were to try to make one, it would be as useful as comedian Steven Wright's full size maps ("Scale: 1 mile = 1 mile"). The best we can do (i.e., should do) is generate a useful list of categories that orients our thinking toward processes and influences that are consequential and leads us to achieve the goals of our assessments.

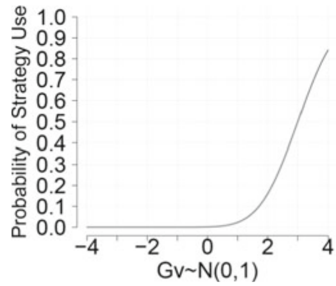
Individualizing Individual Difference Models

One of the central insights of second-wave scholars is that the way in which a person completes a task matters. In some cases, the manner in which the person approaches the test completely alters what is being measured. Although second-wave scholars frequently reiterate this point, rarely do their analyses go further than assertions and anecdotes. However, it is quite possible for these hypotheses to be operationalized and tested empirically (e.g., Hegarty 2010; Kyllonen et al. 1984).

To illustrate, imagine that there is a test that requires the agile use of controlled attention to respond fluently to a sequence of visual images. Such tests are usually thought of as measures of processing speed (Gs). However, unanticipated by the test developers, suppose there is an alternate strategy by which the task can be completed very quickly, even if the person has low Gs. Suppose that this alternate strategy, though not difficult to implement, requires spatial visualization and is typically only discovered by evaluees with high Gv. For evaluees who use the "spatial strategy," Gs is not measured at all, but instead performance mostly depends on Gv. For individuals who do not adopt the spatial strategy (even if their Gv is high), the test is primarily a measure of Gs, as shown in Fig. 21.8.

For the sake of simplicity, suppose that evaluators can easily tell when this spatial strategy is employed, and they check a box when the telltale signs of the strategy's use are observed. Therefore, "spatial strategy" is a binary variable for which 1 and 0 denote the strategy's presence and absence, respectively. As shown in Fig. 21.8, the probability that the spatial strategy will be used is a function of Gv. Only about 1.7 % of people use the spatial strategy. Of strategy users, Gv is, on average, 1.75 standard deviations above the mean. However, even most people with high Gv do not discover the strategy. For example, only 16 % of people with Gv exactly 2 standard deviations above the mean use the strategy, and only 50 % of people with Gv 3 standard deviations above the mean use it. Among non-strategy users, Gv is essentially average (mean = −0.03).





Correlations with Observed Test Scores

Sample	*g*	Gv	Gs
All (100%)	.34	.26	.79
No Strategy (98.3%)	.32	.22	.80
Spatial Strategy (1.7%)	.41	.70	.15

Fig. 21.8 A test of processing speed (Gs) is sometimes a measure of visual-spatial ability (Gv). Note: Φ is the cumulative distribution function of the standard normal distribution

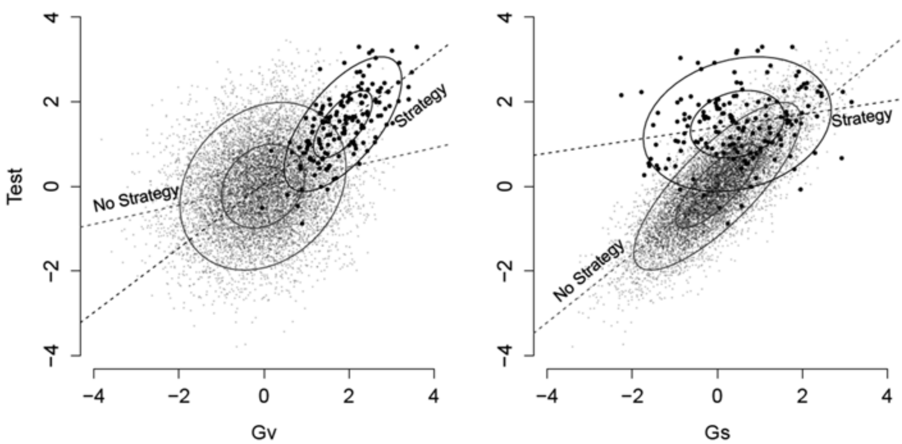


Fig. 21.9 A "Gs" test measures Gv when a spatial strategy is employed

In Fig. 21.8, the loading of Gs on the test is normally 0.8 but is 0 when the spatial strategy is used. Likewise, the loading of Gv on the test is normally 0 but is 0.8 when the spatial strategy is used. Thus, among spatial strategy users, the test is not a measure of Gs but of Gv. Otherwise, the test functions as it was intended. Because typically only people with high Gv discover the strategy, test scores among spatial strategy users are on average 1.4 standard deviations above the mean. In Fig. 21.9, 10,000 cases were simulated according to the model. It can be seen that even individuals with low Gs tended to perform well on the test if the spatial strategy was employed.

What conclusions would come from the typical correlational study in which the effect of the spatial strategy is not modeled? The direct effect of Gv would probably be close to 0 because among non-strategy users the observed correlation between Gv and the "Gs" test is due only to the shared influence of *g*. Clinicians would persist in arguing that this "spatial strategy" phenomenon is real and that it matters. Researchers who only look at the correlational data would counter that Gv's effect is small and thus might treat what clinicians say about the spatial strategy as "unsubstantiated clinical lore."

Many validity-altering test behaviors are impossible to study because they occur exactly once in a practitioner's career (e.g., an extremely bright examinee might score low on similarities because she believes that she can only give single-word responses). Because there is an infinite number of ways to be unusual, practitioners observe rare behaviors regularly. In such cases, there is no need to let the hobgoblin of little minds dictate a consistent by-the-book interpretation of the test score. Unusual behaviors are difficult to interpret precisely because we have no experience with them. It is probably impossible to avoid generating explanations for unusual behavior, but it is generally safer to not take our interpretations too seriously. Testing the same ability with another test is generally the best course of action when the validity of a test has possibly been compromised by unusual test behaviors.

When Will the Second Wave End?

Well-substantiated research findings can never be as rich as the individually tailored case conceptualizations of thoughtful practitioners because practitioners can always take the latest findings and incorporate them into their assessments. What defines second-wave scholars is not a lack of empirical evidence but their focus on individuals. Their insights can be taken seriously and studied rigorously. One of the biggest obstacles to investigating the claims of process-oriented scholars is that researchers from the psychometric tradition tend to rely on latent variable modeling, and only fairly recently has SEM software (e.g., Mplus, OpenMx) made it convenient to evaluate the kinds of if-then relationships that process-oriented scholars tend to highlight.

It is a scientific victory every time that "clinical lore" is found to be substantiated by well-designed research. Indeed, there is reason to celebrate when such lore is found to be an unsubstantiated myth: Sometimes unlearning false facts does more good than learning true ones. However, there will never be a time in which the insights of the second-wave scholars will be obviated fully by scientific findings. The underappreciated genius, William Stern (1900, as cited in Lamiell 2003), said it best:

> [E]very individual is a singularity, a one-time existing being, nowhere else and never before present. To be sure, certain law-like regularities apply to him, certain types are embodied in him, but the individual is not exhausted by these laws and types; there remains ever something more, through which the individual is distinct from others who conform to the same laws and types. And this last kernel of being, which reveals the individual to be thus and so, distinct from all others, is not expressible in the language of scientific concepts, it is unclassifiable, incommensurable. In this sense, the individual is a limiting concept, toward which theoretical investigation strives but can never reach; it is, one could say, the asymptote of science. (pp. 15–16)

We can only add that, as humanists who value the dignity of individuals, our sentiments resonate with the poetic rhythm of Stern's rhetoric; as scientists, we are rather restless to see just how much we can close the gaping chasm between our current understanding of individual cognitive processes and whatever asymptotes there might be.

The upside of the second-wave approach is that it acts as a fail-safe for our incomplete and partially incorrect theories and measures. We do not yet have mature theories and measures that account for the incredibly diverse approaches people take while completing cognitive tasks. As mentioned above, much of the validation efforts for our measurement tools assume a one-size-fits-all interpretation (e.g., "block design measures visualization"). Of course no researcher actually believes that each test means the same thing for everyone, but we do not yet have compelling research that allows us to construct psychometrically sophisticated models that account for countless possible moderators of performance on our tests.

All practitioners have stories about how a standard interpretation of a test score would have resulted in a wildly inaccurate case conceptualization and perhaps an injustice to the evaluee (Schneider 2013a). Second-wave approaches allow us to combine psychometric data and research-based theoretical understandings with our intuition, empathy, and sometimes creativity so that such misunderstandings and injustices are less likely to occur.

The downside of second-wave approaches is that they are only as good as the person using them. Many accounts of the clinical exploits of the masters of these methods have uncomfortable amounts of hagiography and hero worship. It is reasonable to assume that among practitioners there are large individual differences in talent for empathic observation. Second-wave approaches probably magnify the good that can be done by the gifted but also magnify the possible harmful effects of the interpretations and actions of the incompetent. How can we know to which group we belong? As one of Tolstoy's characters in *Anna Karenina* remarked, "No one is satisfied with his fortune, and everyone is satisfied with his wit." By which standards can an individualized interpretation be challenged? Directly challenging another professional is usually considered rude, if not unethical.

Even if we are sincerely open to correction, what feedback mechanisms are available to evaluators who wish to identify and learn from their interpretive errors? Second-wave scholars do have some answers to such questions, but there has never been a compelling account of how ordinary practitioners can use these kinds of approaches to make more valid interpretations than could have been made with more purely psychometric approaches. On balance, however, we estimate that second-wave approaches have done far more good than ill. Nevertheless, there is much room for improvement in second-wave methods and training practices.

Lessons Learned from Wave 3

Just as it is easier for second-wave scholars to speculate than it is for researchers to validate those ideas, it is easier for applied researchers to generate new findings than it is for test developers to incorporate new ideas into well-normed batteries that can be used by practitioners. At the preliminary stages of theory validation, it is comparatively easy to conduct several dozen exploratory studies (with college sophomores by the hundreds) than it is to publish a single instrument with a large nationally representative standardization sample. This situation means that there will always be some amount of pressure to retrofit older tests with new meanings. Theoretical explanations come and go, but digit span is forever!

In the excitement of the early work on cross-battery assessment (XBA; Flanagan and McGrew 1997; Flanagan & Ortiz 2001), a fair amount of such retrofitting necessarily occurred. From the beginning, in communicating with practitioners, care was taken to distinguish between tests placed in a particular category because of strong research findings, weak research findings, or expert consensus. Most early speculations about which tests would load on which factor were later substantiated, but some were not. For example, for decades the Wechsler block design subtest was thought to be primarily a measure of fluid reasoning (Gf). Theory-driven cross-battery factor analyses later revealed that for most people block design primarily measured visual processing (Gv; Flanagan et al. 2013a). Conversely, the Wechsler Arithmetic subtest has been interpreted variously as a measure of working memory capacity (Gsm), quantitative ability (Gq), verbal ability (Gc), and fluid reasoning (Gf), depending on the developmental level of the sample and the diversity of ability measures included in the analyses.

Although much progress has been made in making broad ability classifications, many narrow ability classifications of clinical tests are still untested hypotheses. Part of the reason for this is that the most convincing data from Carroll's work is at the broad ability level, and, as Carroll (1998) warned the field, the taxonomy of narrow factors is not at all settled. Since the publication of Carroll's (1993) masterwork, no one has attempted a systematic reevaluation of the structure of the narrow abilities.

If retrofitting must be done, we have some advice:

1. Be excited about new theories. Use them with gusto! However, remember that they are probably doomed to sound silly a decade or

two from now. It is hard to convey to graduate students how thrilling some of the now-neglected theories of intelligence were in their day. Maintain a healthy sense of history: Most new ideas are not dramatic paradigm shifts, but are actually old ideas with new names and a bit of twist. Often those twists represent real progress. However, some developments are heralded with so much fanfare that it is inevitable that they will ultimately disappoint us.

2. Kaufman and Sattler had the right idea (and still do)—if you must retrofit old tests with new interpretations, all flights of fancy must take off and land on the runway of psychometric constraints. For example, if two tests are allegedly measuring the same thing based on your task analysis, they should probably correlate with each other more highly than they do with other tests.
3. The human mind has trouble grasping nature's preference for continua over dichotomies. As a shorthand for communication, it is common to label a test as belonging to a certain category ("the WISC-IV matrix reasoning subtest measures Gf and the WJ III Visual Closure test measures Gv."). In reality, tests reflect one or more factors, each to varying degrees (e.g., matrix reasoning has a secondary loading on Gv that is slightly larger than visual closure's primary loading.). The important question is not so much "*Does* this test measure factor X?" but "*How well* does this test measure factor X? For whom? Under what conditions?"
4. There are more things in heaven and earth, Horatio, Than are dreamt of in your philosophy. Recognize that the latest psychological constructs are probably not going to be measured cleanly by any tests that predate the new ideas. It is going to be messy, and there is no point in trying to make things fit snugly. For example, Wechsler's picture arrangement subtest is a great test of general ability. Many have attempted to classify it into one theoretical category or another, but none have been terribly convincing.
5. Although classifying tests can be fun, it is a stale enterprise when it becomes an end in itself. It is easy to lose sight of the fact that whether a test measures lexical knowledge, perceptual speed, memory span, and the like cannot be our ultimate concern. What matters is what each test can tell us about individuals. If we do not know what a low score on the WJ III Visual Closure means for the future well-being of an individual, the test's place in a taxonomy is of minor importance.
6. New scientific terms are introduced when cracks and gaps in our understanding are discovered. Some new terms are simply labels for the holes. In contrast, some scientific ideas actually fill in the holes and contribute to soundness of the whole structure. Maybe that is why we call them *constructs*. It is important not to confuse these two types of labels, heeding Melville's warning that mouthing hard words is not the same as understanding hard things (nor is specifying a latent variable model). One sign that something is amiss is when an earwormish bit of jargon jingles and jangles itself into every aspect of a theoretical domain. For example, it is unclear at this point what ability *is not* an "executive function." A construct that is invoked to explain everything is as good as one that explains nothing. An admirable attempt to rescue the term *executive function* from excessive generality can be found in the work of McCloskey and colleagues (2008).
7. When retrofitting old tests with new theories, there is a tendency for one's understanding of an idea to morph and drift over time. In discussions of these constructs, particularly among novices, it is common to hear Ga conflated with Gc, Gsm with Glr, and Gv with Gf and/or Gq. One must read and reread source materials regularly to repair and refresh leaking and lapsing memories.

The Maturation of the Fourth-Wave Approach

Imagine a musical artist who selects the finest musical instruments, the best musicians, and state-of-the-art recording facilities to record a great new song. The music is recorded with the utmost care, and then months are spent meticulously improving the recording with sophisticated postproduction software. All this effort is wasted if people can only listen to the music on old cassette players with distorted tinny speakers. It is likewise a waste of effort for cognitive ability researchers to develop increasingly refined psychometric models if clinicians are not given better tools (both conceptual and technical) with which to make use of those models at the point of contact with individuals.

With powerful and comprehensive theories of intelligence such as CHC theory and PASS theory, well-trained clinicians can describe an individual's cognitive ability profile with much more confidence than was possible in the recent past. Wave 4 models and methods are an advance, but there is still much room for improvement. Flanagan, Ortiz, and Alfonso (2013) present a set of practical guidelines for constructing a battery that will measure broad and specific abilities in a theoretically sound and psychometrically defensible manner. For example, each broad ability is to be measured with two or three qualitatively different tests, preferably from different narrow abilities. Such guidelines will become increasingly nuanced with time; in order to achieve the same level of precision, ability clusters that are less correlated (e.g., visual-spatial processing) need to be measured with more tests than do ability clusters that are more highly correlated (e.g., crystallized intelligence).

In most clinical validation research, the emphasis has been on *which* factor a test measures, not on *how well* it does so. In some cases, the tools and data sets needed to answer how well we measure various abilities are not just within our grasp, but already in our hands. It is possible to use existing data sets and models to calculate how well we are measuring the theoretical constructs we intend to measure. For example, using a model of the WISC-IV developed by Weiss and colleagues (2013), Schneider (2013b) illustrated a method of calculating latent variable scores for individuals along with appropriate confidence intervals around those scores. Such methods, combined with large cross-battery models could be used to give clinicians a far more accurate picture of not only what is being measured but the precision with which inferences can be made. Ideally, instead of rough rules of thumb, clinicians would have guidance about how many tests would be needed to measure a particular ability with a preferred level of precision. Before such information could be available to practitioners, much more cross-battery research will need to be conducted such as the large multi-battery study by Reynolds and colleagues (2013).

This emphasis on precision of measurement can be extended to prediction. For example, it is widely known that phonological awareness deficits are causally related to reading decoding difficulties (Wagner & Torgesen 1987). If we find that a child of otherwise average intelligence has low reading decoding ability and low phonological processing, it is easy to believe that one's work is done—an explanation has been found! Indeed it has…partially. In most cases, though, there is more to the story. In fact, most people with average general ability but with low phonological processing do not read poorly (Schneider 2013a). The vast majority of people with low phonological processing abilities (but average intelligence) actually read at an average level or better!

Though low phonological processing may not be a sufficient explanation for poor reading skills, it is true that it substantially elevates the risk of having reading decoding problems. Therefore, when we encounter a child with reading decoding problems, the question we should ask is not "*Does* low phonological processing account for this child's reading difficulties?" It probably does, at least a little. The better question to ask is, "*How much* of this child's underperformance in reading is explained by low phonological processing?" In what way is this question better? It is actually rare that phonological processing is so low that it is a sufficient explanation of reading problems. If an evaluee's reading underperformance remains mostly unexplained (and it is usually is), we are prompted to search for additional explanations, including other relevant cognitive deficits. Often these additional explanations involve the child's personality, family system, past learning history, current learning environment, and a host of other considerations. A fuller accounting of what caused the child's reading difficulties is likely to lead to substantially better interventions. Focusing solely on phonological awareness deficits might lead to insufficient and ill-fitting interventions that are likely to fail, even though phonological processing really is part of the problem (Schneider 2013a).

Methods for applying complex prediction equations are available, but their use is not yet common, in part because the methods are not fully developed and also because they can be mathematically daunting. A number of scholars are attempting to make it easier to apply complex prediction equations in the assessment of individuals (e.g., Crawford et al. 2012; Flanagan et al. 2013b; Schneider 2010, 2013b).

Wave 5? The Future of Cognitive Ability Test Interpretation

Predictions of the future are often laughable in retrospect, but we will nevertheless hazard a few predictions about how cognitive ability tests will be interpreted in the coming decades. The initial hopes of a successful integration of information processing theories and intelligence theories have not yet been realized (Hunt 2011). Many researchers have proposed various models by which individual differences in cognitive abilities can be seen as parameters of information processing (Kyllonen 2002; Lohman 2000; Schneider & McGrew 2013; Woodcock 1993). As yet, these models are mostly untested hypotheses and are not yet ready to be applied in everyday practice (Floyd & Kranzler 2012). However, ultimately we hope to have a consilient account of the philosophy of mind, neuroscience, universal cognitive processes, and individual differences in intelligence.

One of the barriers to practitioners using the latest theories and methods from cognitive ability research is that cognitive ability theories are now almost exclusively expressed in terms of complex mathematics. Therefore, it is likely that test interpretation, like test scoring, will be increasingly done with the aid of computer software. Instead of loosely applying structural models of cognitive abilities to test interpretation in our heads or using rigid cookbook-style interpretive software, clinicians will increasingly use software to interactively and flexibly apply complex mathematical models directly to individual test data (Schneider 2010, 2013b).

It is likely that cognitive ability test interpretation will be directly incorporated into academic progress monitoring. That is, traditional cognitive ability assessments and the response-to-intervention (RTI) approach will become integrated into a coherent and unified interpretive framework. Complex prediction equations involving cognitive and academic abilities will include important covariates such as past performance, time on task, task persistence, and quality of instruction. This framework is likely to be increasingly informed by dynamic brain imaging techniques and well-developed cognitive information processing models. That is, we will be able to observe information processing deficits and abnormalities in real time as evaluees perform academic tasks. Furthermore, we will be able to monitor directly whether interventions succeed in normalizing the processing deficits we identify.

Like the first automobiles, airplanes, and computers, these innovations initially will be extremely expensive and inequitably distributed and probably will not work very well. If clinicians, researchers, and entrepreneurs are allowed to innovate, these procedures will become cheaper and universally available and, most importantly, will actually work.

References

Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 185–202). New York: Guilford.

Baddeley, A. (2006). Working memory: An overview. In S. J. Pickering (Ed.), *Working memory and education* (pp. 1–31). New York: Elsevier.

Binet, A. (1905). Analyse de C. E. Spearman, the proof and measurement of association between two things and general intelligence objectively determined and measured. *L'Année Psychologique, 11*, 623–624.

Binet, A., & Simon, T. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'annee Psychologique, 12*, 191–244.

Canivez, G. L. (2008). Orthogonal higher order factor structure of the Stanford-Binet Intelligence Scales–for children and adolescents. *School Psychology Quarterly, 23*, 533–541.

Canivez, G. L. (2013). Psychometric versus actuarial interpretation of intelligence and related aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The oxford handbook of child psychological assessments* (pp. 84–112). New York: Oxford University Press.

Carretta, T. R., & Ree, M. J. (2000). General and specific cognitive and psychomotor abilities in personnel selection: The prediction of training and job performance. *International Journal of Selection and Assessment, 8*, 227–236.

Carroll, J. B. (1984). Raymond B. Cattell's contributions to the theory of cognitive abilities. *Multivariate Behavioral Research, 19*, 300–306.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. Cambridge, UK: Cambridge University Press.

Carroll, J. B. (1998). Human cognitive abilities: A critique. In J. J. McArdle & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 5–24). Mahwah: Erlbaum.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38*, 592.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40*, 153–193.

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1–22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action*. New York: Elsevier.

Chapman, P. D. (1988). *Schools as sorters: Lewis M. Terman, applied psychology, and the intelligence testing movement, 1890–1930*. New York: New York University Press.

Cohen, J. (1959). The factorial structure of the WISC at ages 7–6, 10–6, and 13–6. *Journal of Consulting Psychology, 23*, 285–299.

Crawford, J. R., Garthwaite, P. H., Denham, K. A., & Chelune, G. J. (2012). Using regression equations built from summary data in the psychological assessment of the individual case: Extension to multiple regression. *Psychological Assessment, 24*, 801–814. doi:10.1037/a0027699.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Boston: Allyn and Bacon.

Flanagan, D. P., & McGrew, K. S. (1997). A cross-battery approach to assessing and interpreting cognitive abilities: Narrowing the gap between practice and cognitive science. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 314–325). New York: Guilford Press.

Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. Hoboken: Wiley.

Flanagan, D. P., Alfonso, V. C., & Reynolds, M. R. (2013a). Broad and narrow CHC abilities measured and not measured by the Wechsler scales: Moving beyond within-battery factor analysis. *Journal of Psychoeducational Assessment, 31*, 202–223.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013b). *Essentials of cross-battery assessment* (3rd ed.). Hoboken: Wiley.

Flanagan, D. P., Alfonso, V. C., & Dixon, S. G. (2014). Cross-battery approach to the assessment of executive functions. In S. Goldstein & J. Naglieri (Eds.), *Handbook of executive functioning* (pp. 379–409). New York: Springer.

Floyd, R. G., & Kranzler, J. H. (2012). Processing approaches to interpretation of information from cognitive ability tests: A critical review. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 497–523). New York: Guilford Press.

Galton, F. (1888). Co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London, 45*, 135–145.

Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *Journal of Special Education, 40*, 103–114.

Gottfredson, L. (1997). Why g matters: The complexity of everyday life. *Intelligence, 24*, 79–132.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J. P. (1972). Thurstone's primary mental abilities and structure-of-intellect abilities. *Psychological Bulletin, 77*(2), 129–143.

Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. New York: Guilford Press.

Halstead, W. C. (1947). *Brain and intelligence: A quantitative study of the frontal lobes*. Chicago: University of Chicago Press.

Hegarty, M. (2010). Components of spatial intelligence. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 52, pp. 265–297). San Diego: Academic.

Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environment*. Odessa: Psychological Assessment Resources.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence* (pp. 267–300). New York: Wiley.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253–270.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 53–91). New York: Guilford Press.

Horn, J. L. & Blankson, N. (2005) Foundations for better understanding of cognitive abilities. In D. Flanagan & P. Harrison (Eds.) *Contemporary Intellectual Assessment: Theories, tests, and issues* (2nd ed., pp. 41–68). New York, NY: Guilford Press.

Hunt, E. (2011). *Human intelligence*. New York: Cambridge University Press.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of vocational behavior, 29*, 340–362.

Jensen, A. R. (1974). Review of abilities: Their structure, growth and action. *American Journal of Psychology, 87*, 290–296.

Johnson, W., & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence, 33*(4), 393–416.

Kamphaus, R. W., Winsor, A. P., Rowe, E. W., & Kim, S. (2012). A history of intelligence test interpretation. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (3rd ed., pp. 56–70). New York: Guilford.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.

Kaufman, A. S., & Kaufman, N. L. (1983). *KABC: Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R* (p. 1979). New York: Wiley-Interscience.

Kaufman, A. S. (2000). Intelligence tests and school psychology: Predicting the future by studying the past. *Psychology in the Schools, 37*, 7–16.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children-Second Edition (KABC-II)*. Circle Pines: American Guidance Service.

Kyllonen, P. C. (2002). Knowledge, speed strategies, or working memory capacity? A systems perspective. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 415–465). Mahwah: Erlbaum.

Kyllonen, P. C., Lohman, D. F., & Woltz, D. J. (1984). Componential modeling of alternative strategies for performing spatial tasks. *Journal of Educational Psychology, 76*, 1325–1345.

Lamiell, J. T. (2003). *Beyond individual and group differences: Human individuality, scientific psychology, and William Stern's critical personalism*. Thousand Oaks: Sage.

Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 285–340). New York: Cambridge University Press.

Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper & Row.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books.

Major, J. T., Johnson, W., & Deary, I. J. (2012). Comparing models of intelligence in Project TALENT: The VPR model fits better than the CHC and extended Gf–Gc models. *Intelligence, 40*(6), 543–559.

McCloskey, G., Perkins, L. A., & Van Diviner, B. (2008). *Assessment and intervention for executive function difficulties*. New York: Routledge.

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151–179). New York: Guilford.

McGrew, K. S. (2005). The Cattell-Horn-Carroll (CHC) theory of cognitive abilities. Past, present and future. In D. Flanagan, & Harrison (Eds.), Contemporary intellectual assessment. Theories, tests, and issues (2nd ed., pp.136–181). New York.

McGrew, K. S. (2009). Editorial: CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10.

Milberg, W. P., Hebben, N., & Kaplan, E. (1996). The Boston Process Approach to neuropsychological assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed., pp. 58–80). New York: Oxford University Press.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive assessment system*. Itasca: Riverside.

Osgood, C. E. (1963). Psycholinguistics. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 6, pp. 244–316). New York: McGraw-Hill.

Rapaport, D., Gil, M., & Schafer, R. (1945–1946). *Diagnostic psychological testing (2 vols.)*. Chicago: Year Book Medical.

Reynolds, M. R., Keith, T. Z., Flanagan, D. P., & Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology, 51*, 535–555.

Sattler, J. M. (1974). *Assessment of children's intelligence*. Philadelphia: Saunders.

Samelson, F. (1977). World War I intelligence testing and the development of psychology. *Journal of the History of the Behavioral Sciences, 13*, 274–282.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job

performance. *Journal of personality and social psychology, 86*(1), 162–173.
Schneider, W. J. (2010). *The compositator 1.0*. Olympia: WMF Press.
Schneider, W. J. (2013a). Principles of assessment of aptitude and achievement. In D. Saklofske, C. Reynolds, & V. Schwean (Eds.), *Oxford handbook of psychological assessment of children and adolescents* (pp. 286–330). New York: Oxford University Press.
Schneider, W. J. (2013b). What if we took our models seriously? Estimating latent scores in individuals. *Journal of Psychoeducational Assessment, 31*, 186–201.
Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). New York: Guilford.
Schneider, W. J., & McGrew, K. S. (2013). Cognitive performance models: Individual differences in the ability to process information. In S. Ortiz & D. Flanagan (Sec. Eds.), Section 9: Assessment Theory, in B. Irby, G. Brown, & R. Laro-Alecio, S. Jackson (Vol. Eds.), *Handbook of educational theories* (pp. 767–782). Charlotte: Information Age Publishing.
Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology, 15*, 201–293.
Spearman, C. (1915). The measurement of intelligence. *The Eugenics Review, 6*, 312–313.
Spearman, C. E. (1927). *The abilities of man: Their nature and measurement*. London: Macmillan.
Spearman, C. (1930). Autobiography. In C. Murchison (Ed.), *A history of psychology in autobiography* (Vol. 1, pp. 299–334). Worchester: Clark University Press.
Stern, W. (1900). *Über Psychologie der individuellen Differenzen (Ideen zu einer "Differentiellen Psychologie") [On the psychology of individual differences (Toward a "differential psychology")]*. Leipzig: Barth.
Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale* (4th ed.). Chicago: Riverside.
Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41*, 1–32.
Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
Tuddenham, R. D. (1962). The nature and measurement of intelligence. In L. Postman (Ed.), *Psychology in the making: Histories of selected research problems* (pp. 469–525). New York: Knopf.
Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104–132.
Unsworth, N., & McMillan, B. D. (2014). Trial-to-trial fluctuations in attentional state and their relation to intelligence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi:10.1037/a0035544.
Vernon, P. E. (1961). *The structure of human abilities* (2nd ed.). London: Methuen.
Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.
Wainer, H. (1976). Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin, 83*, 213–217.
Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997). Discriminant and predictive validity of the WISC-III ACID profile among children with learning disabilities. *Psychology in the Schools, 34*, 309–319.
Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013). WISC-IV and clinical validation of the four- and five-factor interpretative approaches. *Journal of Psychoeducational Assessment, 31*, 114–131.
Wilks, S. S. (1938). Weighting systems for linear functions of correlated variables when there is no dependent variable. *Psychometrika, 3*, 23–40.
Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment, 8*, 231–258.
Woodcock, R. W. (1993). An information processing view of the Gf-Gc theory. In *Journal of Psychoeducational Assessment Monograph Series: Woodcock-Johnson Psycho-Educational Assessment Battery—Revised* (pp. 80–102). Cordova: Psychoeducational Corporation.
Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock–Johnson psycho-educational battery–revised*. Allen: Developmental Learning Materials.

Intelligence and Culture: History and Assessment

22

Donald H. Saklofske, Fons J.R. van de Vijver, Thomas Oakland, Elias Mpofu, and Lisa A. Suzuki

Human behavior occurs in spatiotemporal and sociohistorical contexts. This chapter discusses intelligence in light of some of its contextual features, including the link between intelligence and culture. The issues are discussed in three parts: historical influences, cultural issues in conceptualizing and assessing intelligence, and intelligence test adaptation and development in the international community. The first section on historical perspectives discusses early conceptions of the assessment of intelligence, including explaining race and ethnic group differences in intelligence scores. The second section provides an overview of contemporary practices, in particular cultural conceptions of intelligence (i.e., how intelligence is defined in different cultures, cultural issues in intelligence assessment) and trends in test use across cultures (e.g., which tests are used in various cultures). The third section highlights issues of equivalence and bias in intelligence tests followed by procedures used to adapt tests or their administration to increase their cultural suitability and issues important to the translation and adaptation of intelligence tests.

Historical Influences in Intelligence Testing

Both the construct of intelligence and its measurement predate the establishment of psychology as a scientific discipline and most likely have a history as long as human civilization. A search through the lexicon of the world's languages both past and present will uncover words and descriptions that capture individual differences in human behavior that reflect intellectual and other cognitive abilities. The overlap as well as the uniqueness in how intelligence has been and is described, defined, observed, measured, valued, and used by cultures, societies, and countries is likely to be of interest to those interested in the concept and measurement of intelligence.

D.H. Saklofske (✉)
Department of Psychology, University of Western Ontario, Social Science Centre (7th floor), London, ON N6A 5C2, Canada
e-mail: dsaklofs@uwo.ca

F.J.R. van de Vijver
Department of Culture Studies, Tilburg University, P. O. Box 90153, 5000 LE Tilburg, The Netherlands
e-mail: Fons.vandeVijver@uvt.nl

T. Oakland
University of Florida, 1921 SW 8th Drive, Gainesville, FL 32601, USA
e-mail: Oakland@ufl.edu

E. Mpofu
Faculty of Health Sciences, University of Sydney, Cumberland Campus Room T-428, Lidcombe, NSW 2141, Australia
e-mail: Elias.mpofu@sydney.edu.au

L.A. Suzuki, Ph.D.
Department of Applied Psychology, New York University, New York, NY 10003, USA
e-mail: Las1@nyu.edu

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_22,

Early Efforts to Define and Measure Intelligence

The first recorded evidence of test development and use occurred in China approximately 3000 years ago. The tests had a decidedly practical purpose, namely, to help identify civil servants. Assessments included essays on civil law, military studies, agriculture, and geography. Subsequent measures of visuospatial perception, creativity, problem solving, and divergent thinking were used somewhat prominently in China. A number of these qualities are consistent with currently identified components of intelligence.

Both the early Roman and Greek civilizations inferred differences among its citizens in what they knew and did. For example, ancient Greek society focused on self-care and community engagement. People were thought to be of normal intelligence if the level of these qualities were similar to their peers and to display diminished intelligence if they were similar to younger persons (Oakland and Harrrison 2008). Aristotle's description of reason or the proposed brain-intelligence connections described by Hippocrates and Plato were forerunners of later thought.

In Europe during the Middle Ages, various "tests" comprised of a few simple questions were used to determine the ability or "mentality" of persons. However, it was not until the nineteenth century that the pioneering work of Sir Francis Galton (1822–1911), Wilhelm Wundt (1832–1920), James McKeen Cattell (1860–1944), and others helped to establish the central role of intelligence within the context of the study of individual differences in psychology, including the effects of race and social class. Their contributions led to worldwide interest in both the concept and measurement of intelligence and to what may be termed "the testing enterprise."

During these early years of psychology, definitions of intelligence remained elusive and gave rise to the development of many tests that later have been discarded as measures of intelligence (e.g., head circumference, visual and auditory discrimination, grip strength). However, the dawn of the twentieth century ushered in new measurement methods based on different theoretical views of intelligence (e.g., knowledge, comprehension, judgment, and reasoning) that were also sensitive to age differences. Alfred Binet (1857–1911), Victor Henri (1872–1940), and Theodore Simon (1873–1961) led this effort after discarding tests of palmistry and head circumference. Binet and Simon's tests, developed between 1904 and 1911, were intended to be used in Paris schools to help identify special needs students. The content of test items was modeled closely on school-like tasks and used for timely selection for remediation on the basis of a learners measured mental age. Empirical evidence that the tests correlated with academic achievement added to their value (Wolf 1973). Interest in ability assessment intensified with the widespread emergence of public education, first in grammar schools (i.e., grades 1–6) and later secondary schools. Their emergence was in response to large increases in the number of students, many of whom came from families that lacked a history of formal education. Educators sought methods that provided information on student's general academic aptitude. The large increase in immigrants into the United States during the nineteenth and early twentieth centuries gave rise to concerns that the country's average level of intelligence was being lowered by their presence and further spurred the development and use of intelligence tests for screening, classification, and selection purposes (Tulsky et al. 2003). In 1916, Lewis Terman (1877–1956) coined the term *intelligence quotient* (IQ), calculated by dividing a person's mental age by his or her chronological age and multiplying this ratio by 100. Subsequent use of ability tests in the military during World War I seemed to secure a lasting place for intelligence testing (and the resulting IQ) in psychology, education, and the public mainstream.

The rapid increase in the number of intelligence tests and their widespread use throughout the twentieth century gave rise to further efforts to develop more comprehensive theories and models that would define the structure of intelligence and in turn guide the measurement and interpretation of intelligence. Early models included those proposed by Spearman (i.e., general and specific factors) and Thurstone (i.e., seven primary mental abilities), and later, during the mid-twentieth century, models from Burt and Vernon (i.e., a hierarchical model that included general, group, specific, and error factors), Jensen (i.e., level 1, memory, and level 2, mental manipulation or *g*), Guilford (i.e., a structure of intellect comprising 120–180 abilities), and Wechsler (i.e., verbal and visual perceptual qualities that contribute to a general IQ). Later twentieth-century models were proposed by Gardner (i.e., multiple intelligences) and Sternberg (i.e., triarchic theory). Cattell and Horn initially proposed two kinds of intelligence, crystallized and fluid intelligence, that later were incorporated into the work by John Carroll, giving rise to CHC theory (named after Cattell, Horn, and Carroll) (i.e., nine broad stratum abilities and more than 70 narrow abilities). Other notable contributions come from Europe (e.g., Das-Luria's PASS model and Piaget's stages of cognitive development) and Israel (Feuerstein's instrumental enrichment model).

While many definitions of intelligence remain in vogue, most include attention to abstract reasoning, problem solving, and the acquisition of knowledge (Snyderman and Rothman 1988). In 1994 in response to controversies generated by the publication of *The Bell Curve: Intelligence and Class Structure in American Life* (Hernstein and Murray 1994), 52 experts on intelligence endorsed a statement published in the *Wall Street Journal* that summarized for public understanding the scientific findings regarding intelligence. Intelligence was defined as "a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson 1994, p. 14). This definition, along with the well-researched CHC model that has influenced many of the tests used in the Western world, has been marketed in the majority world.[1]

[1] We use this term to refer to developing or emerging non-Western countries that constitute the majority of humankind and define community in terms of what it is.

Recognizing Individual and Group Differences in Intelligence

Individual, group, and cultural variability have been apparent from the earliest use of intelligence tests, reflected largely in mean score differences. For example, the US military developed the Army Alpha and Beta tests to facilitate the assessment of recruits who were either literate or lacking in language use and expression, a tradition carried on by Wechsler. As the use of intelligence tests increased in other parts of the world together with an increase in immigration, the issue of observed within- and between-group differences came to the forefront. The major question was whether there were "real" differences between the peoples of different countries, ethnicities, and cultures or whether the observed score differences were a function of the very tests used to assess intelligence. In the mid-twentieth century, Raymond Cattell (1940) developed a "culture-free" (later referred to as a culture-fair) measure of intelligence. In the mid-1970s, Mercer and Lewis (1977) published the System of Multicultural Pluralistic Assessment in response to a need for more culturally fair methods of assessing children from different cultural and ethnic backgrounds. Bracken and McCallum's (1998) Universal Nonverbal Intelligence Test is a recent attempt to minimize test content imbued by culture and language. However, among measures of intelligence, the Wechsler scales are used most frequently to assess both child and adult intelligence.

Test authors and publishers of well-known measures of intelligence along with others in the scientific community have been asked increasingly to provide evidence of the extent to which test content and results provide an accurate assessment when used outside of the context where it was developed and standardized. Even in neighboring countries in which many persons use the same language, such as the United States (US) and Canada, the question of "how well do intelligence tests travel?" was being raised about tests such as the Wechsler intelligence tests that were developed in the United States but used extensively in Canadian schools. This spurred the search for answers to whether intelligence is a universal construct, whether it varies across groups, countries, and cultures, and how to most accurately and fairly measure it.

In the United States, the validity of intelligence tests when used with persons from minority groups, specifically Blacks and Hispanics from lower socioeconomic levels, has been questioned for at least 100 years. The accusation of bias generally is based on an incorrect, outdated model of test bias, one that highlights mean score differences. Modern definitions of test bias do consider differences in mean scores as prima fascia evidence of test bias. Nevertheless, these allegations gave rise to concerns by the public, educational institutions, and psychologists as to whether intelligence tests are fair and unbiased when used with minority groups, especially Blacks and English-language learners (Oakland 1977). Following challenges in the courts of both law and public opinion, the prevailing conclusion among scholars was that intelligence tests are "…not culturally biased against American Blacks or other native born, English speaking peoples in the U.S." and "…members of all racial-ethnic groups can be found at every IQ level" (Gottfredson 1994, p. 14). While group differences in intelligence initially were attributed to genetic and other familial influences, environmental and cultural influences also were recognized as being an important determinant of score variability. The debate regarding the degree to which cultural, genetic, and environmental factors account for racial group differences in intelligence continues (e.g., Rushton and Jensen 2005). Studies, however, continue to support the malleability of intelligence based upon environmental and cultural factors. For example, a study of group differences on the Wechsler tests of intelligence provided a relatively unbiased measure of intelligence for African Americans, Caucasians, and Hispanics in the United States. Demographic qualities, including parents' education and parental expectations for their children and education, accounted for much of the between-group differences (Prifitera et al. 2008; Weiss et al. 2010).

The IQ gap between Blacks and Whites has narrowed somewhat in recent years. For example, IQ differences between Black and White 12-year-olds, in favor of Whites, dropped from 15 points to 9.5 points over the past three decades (Nisbett 2009). Other research found similar results for a broader age group (e.g., Edwards and Oakland 2006; Weiss et al. 2010). Some believe the performance gaps on tests of intelligence for different races depend upon the complexity of the task (Flynn 2012; Jensen 1972, 1973, 1980). For example, subtests with greater *g* loadings (i.e., greater complexity) yield larger group differences. However, more complex test items generally require more implicit and explicit knowledge of the culture of the test (i.e., on tests that have higher cultural loadings) (Helms-Lorenz, Van de Vijver, and Poortinga 2003). As a consequence, the performance difference may be influenced more by knowledge of the culture within which the test was developed than in the complexity of qualities assessed by the test. Moreover, other data indicate scores by Blacks increased on a variety of intelligence subtests (i.e., Wechsler), including those with high cognitive complexity (Weiss et al. 2010).

Although intelligence and culture are linked closely, many aspects of cognitive functioning are universal (Lonner 1980, 2011). For example, most individuals are unlikely to survive if they had not acquired basic cognitive skills (e.g., Piagetian conservation skills). However, results from cross-cultural research on intellectual assessment also make clear that these basic skills often are difficult to identify, notably in non-Western cultures. Thus, interpretations of cross-cultural differences in intellectual skills remain controversial and present complex challenges to test developers, researchers, and practicing psychologists. Genetic and cultural differences together with assessment issues cannot be overlooked.

Cultural Issues in Conceptualizing and Assessing Intelligence

Contemporary Practices

Despite spectacular advances in assessment technology within Western societies, recognition of intellectual abilities that matter to the majority non-Western world and ways to measure them remain underdeveloped (Mpofu and Ortiz 2009; Serpell and Haynes 2004). A reason for this lag in development is due, in part, to the relatively limited research investment in understanding the nature and types of abilities that are valued in these settings. Additionally, there has been a long-standing belief that concepts of intelligence that emanate from Western views are transportable to non-Western settings, sometimes referred to as an “etic” perspective. While this view underlies efforts to discover and map universal traits in intellectual abilities, there is no cross-cultural consensus about the nature of human intellectual abilities and how these should be recognized or measured (Marfo 2011; Mpofu and Nyanungo 1998; Pence 2011; Serpell 2011). Thus, the results of studies that examine the transportability of Western developed tests to emerging countries indicate their limited incremental value and that their results often merely highlight mean score differences between Western and non-Western persons rather than shed light on the nature of intellectual abilities of persons in non-Western cultures. This latter emic perspective is illustrated in the following sections.

Cultural Values and Human Abilities

Cultural values play a significant role in the ways human abilities are recognized, displayed in context, and measured. The meanings that a culture assigns to displayed abilities and other behaviors influence the behaviors that define intelligence in those settings. An understanding of the important intellectual abilities in a cultural setting requires the framing of the abilities from the perspective of those who display and find meaning in a particular context (Nsamenang 2006; Mpofu et al. in press; Serpell 2008, 2011). This includes learning and knowing the cultural contexts that define those abilities and the typical indicators of them.

Concepts of intelligence in the non-Western world are represented best by implicit theories about abilities that the culture prioritizes (Sternberg 1985). Implicit theories about behaviors are those that people use to guide their conduct in a broad range of situations, typically representing shared underlying meanings and social judgments they make on their and others’ behavior. An understanding of the implicit theories of ability held by those engaged in a culture provides a productive way to begin the journey toward understanding and appreciating the diversity in human abilities that are valued.

Knowledge of implicit theories of intelligence held by a cultural community is prerequisite to constructing explicit (or formal) theories for the design and use of appropriate assessments for such abilities. As an example, Sternberg et al. (2001) observed tacit abilities to be highly valued in a Kenyan indigenous Luo cultural-linguistic community. On the basis of that understanding, the authors developed the test of tacit knowledge for natural herbal medicines for use with Luo community children to assess their knowledge of common illnesses and standard herbal treatments for those illnesses in that community. The test is designed to access children’s acquired knowledge about the illnesses and herbal treatment regimens through their experience and informal observation. The test presentation format uses 22 stories that require children to identify an illness that a story depicted and the appropriate herbal treatment for it. Storytelling and interpretations are valued procedures to tap into tacit knowledge held by those living in sub-Saharan African communities.

Structure and Function of Intelligence. An examination of both ancient and modern views from across the globe immediately shows that there is not a universal consensus about what intelligence is, what causes it, and how to best measure it. Past and contemporary views of intelligence in China reflect some major differences from Western perspectives (Yan et al. 2009). For example, key positive human qualities are more akin to linking Western descriptions of intelligence with personality and conative factors (Yan and Saklofske 2004). Descriptions of intelligence in majority world settings may be less about the “what” (structure) than the “how” (function) of abilities, and of course, what a culture values. This may be explained, in part, by prioritizing lived (and thus a personal understanding of practical-oriented) intelligence rather than its hypothesized conceptual structure. For example, an investigation of concepts of intelligence among the Luo of Kenya identified four terms that referred to intelligence: *rieko*, *luoro*, *winjo*, and *paro* (Grigorenko et al. 2001). *Rieko* refers to the context- and task-specific knowledge or skill; thus, school-related *rieko* refers to ability to perform school-related tasks. *Rieko mzungu* refers to competence in white man's (mzungu) technology. A person's *luoro* is expressed by his or her ability to display social responsibility by being considerate of others, observant of appropriate social etiquette, and willing to share resources. *Winjo* refers to behaviors that display appropriate deference to adults, the elderly, and authority figures. *Paro* refers to one's ability to be innovative and creative and to complete tasks in a timely fashion. Intellectual abilities underlying *luoro* and *winjo* would be evident in the quality of a person's social participation, whereas those for *reiko* and *paro* would be evident from the efficiency in completing tasks. Some studies (e.g., Grigorenko et al. 2001; Serpell 2008, 2011) reported performance ratings on contrived Western measures of intellectual ability that poorly predict success on culturally embedded intelligence measures. In non-Western culture settings, higher levels of intelligence often are aligned with practical or functional intelligence (Sternberg 1999) together with social emotional and sensitivity (Mayer et al. 2008).

Qualities of Intelligence in Majority Culture Settings. Six qualities seemingly define the intellectual functioning in non-Western, majority world contexts: interpersonal relationships, socially oriented planning, decision-making and problem solving, socially responsible resource management and utilization, cultural historicity and acknowledging highly regarded others who took similar actions, and social worth and productivity (Dasen 2011; Mpofu 2004; Mpofu et al. 2007; Mpofu et al. in press). In these settings, the activities, domains, and contexts for the valid appraisal of intellectual competence characterize individual and group participation in culturally significant daily activities. For example, a person's intelligence could be described on the basis of his or her abilities to achieve practical outcomes for oneself and important others (e.g., family or clan). To illustrate, Ugandan Baganda and Batoro villagers regard intelligence as socially oriented behavior that benefits the social collective (Wober 1974). This view is consistent with current concepts of social and emotional intelligence even if various cognitive capacities may be underlying factors.

Observing the social responsiveness of actions, processes, and outcomes across cultures provides further indications that intelligence is viewed differently from current Western views that are influenced heavily by psychology and captured in the tests used to assess intelligence. In many non-Western cultures, the timeliness of a selected course of action (rather than the speed of action), its social or collective intent (protecting the honor and social respect of others by sharing the credit for a job well done with them), and recognition of history and traditions are qualities that define those with high intelligence. All persons are presumed to be intellectually competent unless there is compelling and consistent evidence to the contrary. This allows for a high latitude of "ability appreciation" and a discounting process that can be applied whereby behaviors that are less than expected are judged and explained by others in light of the context in which the behavior occurred rather than as a reflection of the person's ability. Measures of wisdom and the ability to associate appropriate social worth to actions as well as to decisions of one's self and others (Sternberg and Jordan 2005) would appear to be more akin to intelligence in these contexts.

Intelligence Tests and Testing

The above discussion emphasizes that, while intelligence is widely considered a cornerstone in describing individual and groups differences (and similarities), how intelligence is defined and assessed varies across time and cultures.

While such diversity is both acknowledged and respected, the Western perspective on intelligence appears to be most widespread and growing. The dominance of the Western perspective in studies and models of intelligence has strong historical and intellectual roots as well as a strong economical component. Western developed intelligence tests are used on a massive scale, not just in Western countries. Therefore, developments in international testing constitute another place where culture and intelligence meet.

Setting the Stage. The proliferation of measures to assess intelligence and the importance of tests and testing in psychological science and practice warrant further discussion before examining intelligence tests in the context of culture. Test use constitutes the dominant method of measurement in psychology. While varying methods of measuring (e.g., observation, interview) the plethora of identified psychological factors (e.g., extraversion, anxiety, resilience, self-concept, and emotional intelligence), standardized tests are likely to be the most common methods and, as discussed above, have their roots in earlier tests developed by Galton, Cattell, and Binet. A test consists of a standard set of questions or items designed to assess knowledge, skills, interests, or other characteristics of an examinee (VandenBos 2007). Minimally, tests should display key psychometric qualities, including reliability and validity, and result in providing practical ways to assess desired constructs (see Saklofske et al. 2013). Knowledge that "psychological test validity is strong and compelling" and that "psychological test validity is comparable to medical test validity" is reassuring (Meyer et al. 2001, p. 128). Many psychologists would contend that psychological tests constitute the flagship of applied psychology with standardized intelligence testing as its most widely used invention.

Earlier in our history, most psychological tests were developed in universities to support research. Although this practice continues, it now is overshadowed by testing companies working in partnership with test authors to develop tests for clinical and other applied markets. Test use is the strongest in the applied and practice areas of psychology, especially school and clinical psychology, as well as industrial/organizational psychology especially related to human resource development.

The steady development of testing resources during the first third of the twentieth century warranted the emergence of the *Mental Measurements Yearbook* that both catalogued and reviewed commercially available standardized tests. Later, *Tests in Print* emerged, thus allowing the *Mental Measurements Yearbook* to focus on test reviews. It lists virtually all English-language standardized tests. Information in *Tests in Print* can provide some insights into market-driven needs. The 3,003 tests referenced in the eighth edition published in 2011 are divided into 20 categories, thus providing a partial summary of market-driven forces that help identify tests that may be used and thus what concepts command the greatest public and professional interest (Murphy et al. 2011). Among these tests, 21 % assess personality, 20 % assess vocations, and 7 % assess intelligence. Some concepts of personality include intelligence as an important component. Moreover, psychological assessment as practiced by clinical, school, industrial/organizational, and counseling psychologists as well as neuropsychologists may include an assessment of intelligence. Thus, this information suggests applications of the concept of intelligence through test use are important to psychologists.

Test Use Internationally. International surveys of test development and use with persons of all ages have not been conducted, although recent publications such as *Pruebas Publicadas en Española: An Index of Spanish Tests in Print* provide an indication of the growth of tests and testing in other countries. However, our understanding of the international status of test development and use with persons from birth to death is dependent somewhat on anecdotal evidence. Besides psychologists, various other professions use measures of cognitive abilities including psychiatrists, educators, counselors, and speech pathologists. Attention to global test use also has been addressed by the International Test Commission, in part through its Guidelines for Translating and Adapting Tests (see www.intestcom.org for the 2010 updated version) in terms of context, test development and adaptation, administration, and documentation/score interpretation.

Test Use Internationally with Children and Youth. Two surveys of test development and use with children and youth provide some indications of international test use, including measures of intelligence, with children and youth (Table 22.1). The 1989 survey acquired information from school psychologists in 44 countries (Hu and Oakland 1991; Oakland and Hu 1992). The 2012 survey acquired information from school psychologists in 77 countries.

Summary of the 1989 Survey Data. Of the ten tests identified in terms of popular usage, five were measures of intelligence (Table 22.1) while the remaining tests assessed visuomotor skills, personality, and psychopathology. Additionally some countries reported using the Frostig Developmental Test of Visual Perception to provide a culture-fair assessment of intelligence. Thus, at that time, school psychologists in international settings used intelligence tests somewhat prominently.

Summary of the 2012 Survey Data. Respondents to the more recent survey identified 606 tests. Their rank ordered areas of focus were as follows: intelligence, achievement, language, personality, visuomotor qualities, school readiness, social and emotional qualities, motor skills, and adaptive behavior. Again, among the top ten tests cited among the 77 reporting countries, six measure intelligence. The remaining tests assessed visuomotor skills, psychopathology, achievement, and personality.

When asked to indicate the purposes of using tests, the respondents identified diagnosis as being most prominent followed by attempts to acquire a better understanding of children's achievement, make admissions decisions, support guidance services, estimate academic aptitude, and assist in vocational selection. Test use by school psychologists is most common among persons ages 6–12 years followed by ages 13–18, 3–5, and over 18 years; as expected, tests were least often used by school psychologists in the assessment of infants.

Table 22.1 Tests used with children and youth

Top ten tests in 1990 check date, is it 88, 89 or 90???
Wechsler intelligence scales for children
Ravens Progressive Matrices
Bender Gestalt Test
Rorschach Inkblot Test
Stanford-Binet intelligence scales
Wechsler adult intelligence scales
Thematic Apperception Test
Differential Aptitude Test
Minnesota Multiphasic Personality Inventory
Frostig Developmental Test of Visual Perception
Top ten tests in 2012
Wechsler intelligence scales for children
Ravens Progressive Matrices
Bender Gestalt Test
Wechsler Preschool and Primary Scale of Intelligence
Kaufman Assessment Battery for Children
Child Behavior Checklist
Wide Range Achievement Test
Children's Memory Scale
Children's Apperception Test
Wechsler Individual Achievement Test

Implications of the Survey Data. Assuming that test use is market driven, we could conclude the assessment of intelligence constitutes a steadfast public interest and remains a common feature of the work of psychologists internationally, especially school psychologists. During the 23 years covered by these survey data, the Wechsler scales for assessing children's intelligence, the Raven's Progressive Matrices, and the Bender Gestalt Test have maintained their prominent ranking. The Wechsler scales are used in all reporting countries. However, 74 countries use older versions of this test than the current WISC-IV. The WISC-V is projected to be published in 2014. The more recent versions of the Wechsler Preschool and Primary Scale of Intelligence and the Kaufman Assessment Battery for Children replaced the previously revered Stanford-Binet as well as the Frostig measure. Two projective-type measures of personality (e.g., Rorschach Inkblot Test and the Thematic Apperception Test) along with the Minnesota Multiphasic Personality Test dropped off the earlier list in lieu of greater use of the measures of children's psychopathology, memory, and achievement seen in the 2012 data.

In much of the Western world, the CHC model of intelligence (McGrew and Flanagan 1998) provides one of the leading descriptions of intelligence and often is used to guide the development or revision of commonly used tests such as the Woodcock-Johnson III Tests of Cognitive Abilities—a measure that is being used increasingly in the United States even though it is not included in the 2012 listing. However, none of the tests listed in Table 22.1 are explicitly grounded in CHC theory. Thus, current tests seemingly underrepresent the complexity of the concept of intelligence, at least as reflected by the CHC model. However, clinicians can use a cross-battery approach to use various measures and then to recast their scores to be more consistent with this model (Flanagan et al. 2013). The Wechsler scales assess five major factors that line up with CHC model without trying to replicate it (Weiss et al. 2013a, b; see *Journal of Psychoeducational Assessment* Special Issue, 2013).

One may wonder why intelligence may be valued so highly and measured so widely compared to measures that assess other important psychological constructs and factors. Although we are not suggesting that measures of intelligence are overused or do not provide important or relevant information, we acknowledge other rich and important testing resources that, if and when used properly, could enhance our understanding and promotion of child growth and development as well as address issues important for understanding infants and adults. Furthermore, with the exception of Mpofu's important and pioneering teamwork in Botswana (Mpofu et al. 2012; Mpofu et al. in press), ultimate consumers of test results (e.g., children and youth, their parents and teachers) rarely are asked to identify the information they would like to obtain through tests. Instead, the transporting of tests from developed to the majority world likely occurs through the efforts of psychologists, often trained abroad, who return to their developing country and attempt to use or even replicate some of the major tests found commonly in the practices of the more developed countries—principally those developed in the United States.

Test-related conditions often differ in developed and emerging countries. Emerging countries tend to use foreign-developed tests that are older versions, have outdated norms, and are not standardized for use in their countries. They also are more likely to use fewer tests for multiple purposes. For example, in one country, the Wechsler scale is used to assess children's school readiness, intelligence, academic achievement, speech/language, social/emotional development, adaptive behavior, motor development, and neuropsychological qualities. School psychologists in emerging countries are likely to use fewer tests that assess personality measures and, if used, to rely on interpretations gleaned from theory, not from the use of norms.

Conditions for test development differ between developed and developing countries. In general, test use is strongest in Australia, Canada, most Western European countries, and the United States. These countries display strong beliefs in science, technology, individual differences, and meritocracy (Oakland 2009). In contrast, both test development and use are lower in Asia, Eastern Europe, and Latin America, among the 22 Arab countries and, with the exception of South Africa, lowest in the 54 African countries. Compared to developed countries, those in the majority world may have smaller populations and thus smaller test markets, fewer test development companies, thus less financial support for test development, together with lower professional standards for those who use tests. They also have fewer university graduate programs that develop skills important to test development, fewer scholarly journals that feature assessment and test-related issues, and fewer professional associations that advocate for high legal and ethical standards for test development and use.

Intelligence Test Adaptation and Development

Equivalence and Bias in Intelligence Testing

Based on the discussion above, one may presume that intelligence is understood best in the context of culture and the identification of "universals" across groups (e.g., common factors such as problem solving, reasoning ability). Thus, we should strive to accurately measure and understand universal traits within the cultures in which they are displayed and whether the traits are culturally meaningful and important. For example, the factor structure of the WISC-III was found to be remarkably robust across more than a dozen developed countries representing Asia, Europe, and North America following careful translation of test items or modifications of content defining the factors and subtests (Georgas et al. 2003). Other evidence suggests that the structure of intelligence is invariant across cultures (Van de Vijver 1997). However, as we can infer from the above discussion, instruments developed for and validated in and for one cultural context may not retain their meaning when used in other cultural contexts; their meaning and usefulness need to be demonstrated empirically. While we agree that how intelligence is viewed, manifest, valued, and assessed within and between cultures may vary, we also contend that there are "universal" or common aspects that can be described and in turn used in a sensible, ethical, and culturally sensitive way with and for the benefit of the individual and larger group. The discussion to follow outlines the rigorous methodology needed to achieve this outcome.

Understanding equivalence is imperative in the assessment of intelligence as it addresses whether the procedures and formats of the measure, results, and interpretations are equally familiar and relevant across diverse populations. Cultural equivalence refers to whether "…interpretations of psychological measurements, assessments, and observations are similar if not equal across different ethnocultural populations" (Trimble 2010, p. 316). Cultural equivalence, a higher-order form of equivalence, is dependent on measures that meet specific criteria that indicate a measure is appropriate for possible use with other cultural groups beyond the one for which it originally was developed. Upwards of 50 or more types of equivalence may impact interpretive and procedural practices in order to establish cultural equivalence (Trimble 2010).

Various forms of equivalences may need to be known in order to use a measure of intelligence cross-nationally with confidence (Butcher and Han 1996; Helms 1992; Lonner 1985; van de Vijver 2001). These include (a) functional (i.e., whether the construct of IQ as operationally defined occurs with equal frequency across groups), (b) conceptual (i.e., item information is familiar across groups and the construct of intelligence means the same thing in various cultures), (c) linguistic (i.e., language used has similar meaning across groups), (d) metric (i.e., a scale measures the same behavioral qualities or characteristics and the measure has similar psychometric properties in different cultures), and (e) scalar (i.e., score differences reflect the same degree, intensity, or magnitude for different cultural groups). Scalar equivalence is established only when mean differences in scores can be attributed to genuine cultural differences and not to any form of bias (Butcher and Han 1996). Thus, evidence is needed that the measure is operating appropriately in various cultural contexts. Most measures of intelligence and indeed most psychological measures have not attained scalar equivalence across diverse cultural groups. Thus, the use of measures of intelligence and the meaning of racial and ethnic group differences in scores obtained on these measures remain somewhat controversial.

Test developers address issues of equivalence through various procedures. They include expert panel reviews (i.e., professionals review item content and provide informed judgments regarding potential biases), examination of differential item functioning (DIF) between groups, statistical procedures that compare psychometric features of the test (e.g., reliability coefficients) based upon different population samples, exploratory and confirmatory factor analysis and structural equation modeling (i.e., examination of the similarities and differences of the constructs structure), and mean score differences while taking into consideration the spread of scores within particular racial and ethnic groups as well as between groups.

State-of-the-art measures of intelligence may use expert review panels, racial and ethnic oversampling, and inclusion of statistical practices that address multiple forms of equivalence through the use of various statistical procedures. Despite sophisticated efforts, problems may remain. For example, the Wechsler scales have been translated and renormed in several countries around the world (e.g., Georgas et al. 2003). The reliability and validity of the measures have been examined in their new cultural contexts, and statistical procedures were conducted to establish linguistic, structural, and conceptual forms of equivalence. However, concerns continue to arise as clinicians struggle to apply the measure in diverse cultural contexts. For example, even with procedures deemed appropriate to establish linguistic equivalence (van de Vijver 2001), a Spanish version of the measure developed in a Spanish-speaking country may not be suitable for use with Spanish-speaking individuals in the United States. Language alone poses a significant obstacle given vast number of languages currently spoken in the United States. There are currently 6,909 languages spoken worldwide and 381 coded in the US Census (Ryan 2013).

Bias in Intelligence Tests. Results from cross-cultural studies of intelligence raised concern about the suitability of such tests when used with persons in a different and unfamiliar culture or with those unfamiliar with the source test's language. The Army Beta was the first test to remove the effects of language on the assessment of intelligence. As noted above, Cattell (1940) first attempted to develop a measure free from cultural bias. Cattell's seminal work on culture-free intelligence tests has led to various refinements that interestingly became less ambitious over time. More recently, Jensen (1980) proposed that, although culture-free tests are unattainable, some instruments reduce the impact of culture (i.e., culture-reduced tests). Ironically, they do not differ appreciably from Cattell's original measure. Importantly, Jensen proposed to combine adequate test design with strict psychometric criteria for assessing the cultural appropriateness of instruments. In contrast, Cattell seemingly believed that careful test design alone would be able to resolve the problem of cultural bias in psychological assessment.

During the last decades, bias in intelligence tests has been addressed from three perspectives: conceptual advances in theories of intelligence, improved and more abundant statistical methods, and conceptual advances in test adaptations. The CHC model of intelligence has gained wide acceptance in intelligence research, including cross-cultural intelligence research. Its acceptance is fairly widespread (Carroll 1993; McGrew 2005), and the model has demonstrated a robustness that seems to show it is applicable universally (Berry et al. 2011; Van de Vijver 1997). Evidence for this universality comes from numerous cross-cultural studies in which factors underlying intelligence batteries were compared. Although most research underlying this model was conducted before the CHC theory was formulated, much evidence suggests that its fundamental features of intelligence (e.g., fluid and crystallized intelligence) can be found in all cultures. However, this universality should not be taken to imply that all measures of intelligence work equally well in all cultures.

Improved statistical methods are available for use in cross-cultural research. For example, structural equation modeling has become popular. Considerable toolsets exist to examine whether items, subtests, and tests can be compared across cultures. Important conceptual advances in test adaptations are reviewed below in considerable detail. These developments, taken together, indicate that much progress has been made since Cattell formulated the problem of cultural bias in intelligence. In addition, few cross-cultural researchers believe that our best intelligence measures are culture-free. However, researchers often mistakenly believe that adequate test design and psychometric evidence on intelligence tests examined cross-culturally and cross-nationally alone are crucial to attempts to understand and reduce cultural bias (Mpofu and Ortiz 2009).

The term *bias* is the generic term to explain cross-cultural differences (Van de Vijver 2003; Van de Vijver and Leung 1997). The term refers to a problem common in cross-cultural studies that data from tests of intelligence and other traits obtained from different groups may not reflect the groups' traits accurately. If scores are biased, their psychological meaning is unique to them (i.e., group dependent). Group differences in assessment outcomes may be accounted for, at least to some extent, by other psychological constructs or measurement artifacts (e.g., stimulus familiarity). Bias is not an intrinsic property of an instrument but arises when a particular instrument is used to compare an individual to a larger group or make between-group comparisons. An instrument that works well in two groups may not work well in a third group. Some instruments are more susceptible than others to bias. For example, an examination of the Georgas et al. (2003) cross-cultural study of the WISC-III showed that the greatest number of changes were made in the subtests that comprised the verbal comprehension factor. In contrast, the Raven's Progressive Matrices often are considered to be among the least culturally biased measures of intelligence and one of the highest in terms of *g* loading.

Sources of Bias. There are three possible bias sources in cross-cultural research: construct bias, method bias, and item bias.

Construct bias occurs when the construct measured is not identical across groups. For example, this may occur when data suggest the very basis of intelligence differs for people from, say, the People's Republic of China and Zimbabwe or Canada and Paraguay. This relativistic viewpoint lacks evidence. Another more likely example occurs when data from an intelligence test administered in two countries display different factor structures. Similarity of factor structures provides evidence for the identity of underlying structures. Thus, when factor structures of test data obtained in different cultures differ, one may conclude that the cognitive processes and structures may not be universal or that problems exist with the cultural suitability of one or more subtests.

The history of cross-cultural psychology is replete with examples. For example, Porteus (1937) administered his maze test, a paper-and-pencil test in which the examinee must trace a way out of a schematically drawn maze, to various cultural groups, including the Khoisan in the Kalahari Desert. His participants were not used to working with drawings on paper had some difficulty completing the tasks. Porteus concluded that their spatial skills were poorly developed, not realizing that his conclusion contradicted by their display of the excellent tracking skills in the relatively cueless Kalahari desert.

Method bias can result from sample incomparability, instrument characteristics, tester effects, and the method (mode) of administration. Stimulus familiarity (i.e., knowing the materials used in the test such as pictures) provides a common example in intelligence tests. Tests administered to children from different cultures are unlikely to have stimuli that are equally familiar to all children. For example, Cattell originally thought that basic geometric shapes (e.g., squares, triangles, and lines) would be equally familiar to children in all cultures. This assumption is incorrect. For example, a child's early social and environmental contexts can influence exposure to geometric shapes. Although some problems associated with the lack of familiarity may be overcome by test instructions or providing additional examples prior to commencing testing, one should not assume that pervasive cultural differences in antecedent experiences can be undone through these methods.

Item bias or differential item functioning can assist in detecting bias due to item level anomalies (e.g., inadequate item translations). An item is biased if persons with the same standing on the underlying construct (i.e., they are equally intelligent) who come from different cultural groups do not display the same expected score on an item. If a geography test administered to pupils in Germany and Australia contains the item "What is the capital of Germany?" the item would be biased in favor of German students as they can be expected to display higher scores on the item than Australian students when pupils with the same total test score are compared.

Construct and item bias has been studied most frequently. The reason for this interest is readily understandable. Unlike method bias that usually can be assessed only by adding items or subtests to an existing test battery, the statistical analysis of construct and item bias is more straightforward and does not require additional data. Thus, method bias often is a neglected source of bias in cross-cultural test use. Very few cross-cultural studies address confounding sample differences. For example, such confounding differences may include previous test exposure, educational quality and opportunity, and familiarity with the test's tasks.

Translating and Adapting Intelligence Tests

Adaptation, Adoption, and Assembly. Conceptual advances in test translation have shifted significantly during the last decades. Early test translations focused largely on changing a test's language from the original source test to a target test. Decades later, the importance of examining the quality of the translations became more apparent. The introduction of back translation procedures marked an important hallmark (Brislin 1970). This procedure translates an instrument from its source language to a target language and then back from the target test to the original source language (i.e., a forward and then a backward translation). Then, the original instrument and the back-translated version are compared. Similarity between these two versions is seen as support for the quality of the translation. This procedure works well when all items can be translated closely in relation to language content and meaning (e.g., simple biographical questions can be checked adequately using this procedure). However, the procedure does not work well when an instrument has idiomatic expressions or has other linguistic features that are difficult to translate (e.g., words such as *him* or *her* are not used in the same way across languages).

Dissatisfaction with translation procedures led to the recognition that a test translation is more than a linguistic exercise. Adapting a test for use in another culture that uses a different language requires a combination of linguistic, cultural, and psychometric expertise. Thus, while linguistic skills remain important, experts with cultural and psychometric knowledge also are needed. For example, the item "Name the president of the United States" may not be difficult to translate into other languages. However, we first need to know the purpose of the item. Is it to test the examinee's knowledge of a specific person or is it intended to assess the name of one of the best-known persons in a country? If the purpose is the latter, then an equivalent item in French could ask examinees to name the French president. If this latter item were used in France, one needs to determine whether the American and French presidents are equally known in their countries. The psychometrically important question is whether the item will perform in the same manner in the two languages. This example is intended to illustrate that translating a test often requires a combination of professional skills, often provided by a group of persons, not a single person.

Thus, the translation of an item (or test) can use three methods: adoption, adaptation, and assembly. A "close" translation may yield an adequate version in the target language (i.e., an adoption method; Van de Vijver 2003; Van de Vijver and Poortinga 2005). Secondly, an item (or test) may need some modification to make it suitable in the target language (i.e., an adaptation method). For example, the term *dollars* should be converted to *euros* and *miles* to *kilometers*. Thirdly, an item may be entirely inappropriate in the target culture. For example, a Dutch intelligence test has an item about bacon. This item is inappropriate for Muslim immigrant children in the Netherlands, given the food taboo on pork meat in this group. In such a case, an entirely different item needs to be constructed (i.e., an assembly method). Test translations typically have relied on adoption methods. The term *adaptation* (e.g., changing parts of the stimulus or instrument) currently is a generic term used for translations. Assembly methods lead to more significant changes and may result in constructing a new instrument. The above example about the maze test may illustrate the use of this method. The measurement of special skills among Khoisan through a paper-and-pencil version of mazes test is not appropriate as it has such low ecological validity. Therefore, a completely different instrument needs to be constructed to measure their spatial skills.

Thus, although close translation methods have been relied on somewhat exclusively, current efforts focus more on preparing test instruments that are suitable in their new cultural context. Adaptations and assembly methods are preferred when close translation methods are unsuitable.

Types of Adaptation. The need for adaptations can arise from various sources (Malda et al. 2008; Van de Vijver and Leung 2011). Cognitive tests can be adapted in five ways: construct driven, language driven, culture driven, theory driven, and familiarity/recognizability driven.

Construct-driven adaptations address differences in definitions of psychological concepts across cultures. For example, there are various non-Western studies of everyday conceptions of children's intelligence in which obedience and rule compliance are components of intelligence (Goodnow 1984; Hess et al. 1980). A full assessment of intelligence in these cultures would require the assessment of obedience.

Language-driven adaptations address the unavailability of synonyms across languages. For example, a person who displays vigorous energy in an uncivil manner may be called aggressive in English (e.g., an aggressive salesperson). Translations in other languages presumably may amount to a description of the concept rather than a direct translation in order to avoid the connotation of aggression with violence (e.g., the translation could become "a forceful and energetic salesperson"). Other adaptations could be a consequence of language differences using personal references. For example, the English word *friend* can indicate both a male and a female person, while other languages use different words for a male and a female friend.

Culture-driven adaptations address different cultural norms, values, communication styles, customs, or practices. For example, the translation of the English *you* requires cultural knowledge about appropriate modes of address in tests when languages distinguish formal and informal addresses (e.g., *vous* and *tu* in French). The distinction can be simple yet important. For example, Dutch also draws a distinction between these two modes, but the conditions in which the two forms are used are not entirely identical in French and Dutch. Thus, one would use the informal address in a Dutch test for students and a formal address in a test for adults where the French version would use the formal address for both populations.

Theory-driven adaptations address changes based on underlying concept or theory. For example, digit span items should have digit names that are all of similar length. However, similarity in digit length may be lost when the items are translated into another language.

Familiarity/recognizability-driven adaptations address differences in the familiarity with task or item characteristics. For example, drawing cords on phones is inappropriate for those who are exposed only to mobile phones.

An Illustration of Test Adaptation: The KABC-II. As part of a large study of the influence of micronutrients on children's intellectual skills, a development team (Malda et al. 2008, 2010) adapted the Kaufman Assessment Battery for Children, Second Edition (KABC-II; Kaufman and Kaufman 2004) for use with urban children in Bangalore, India. The children attended primary school and came from poor families (e.g., average monthly income was slightly above $56 USD). Many adults were illiterate or had only a few years of education. Most houses had one or two rooms, and the average number of people in a household was 5.8. The homes had few or no toys and usually no learning materials other than school books. Most families owned a television. Children either played in the streets or watched television when not doing chores.

The KABC-II (a revised and re-standardized second edition of the KABC) is an individually administered measure of intelligence designed for US children and youth ages 3 through 18.

It measures short-term memory, visual processing, long-term storage and retrieval, fluid reasoning, and crystallized abilities.

Locally developed test were not available for Kannada-speaking children. The choice of the KABC-II was based on several criteria. The instrument has some attractive features for use in a non-Western context. For example, the instrument has a firm theoretical basis in the Luria neuropsychological model but also one that is compatible with the CHC model (Carroll 1993; McGrew 2005). Both the Luria and the CHC models were "assumed" to be valid in this population. The test administration also is less dependent than other tests on the child's use of receptive and expressive language skills. Additionally, sample items are provided prior to administering the test to help ensure a child understands the task demands before commencing the assessment.

The research team was aware that the work may involve five types of adaptations. The team attempted to avoid various sources of bias by selecting the KABC-II. For example, the theoretical underpinning of the KABC-II makes it unlikely that construct-driven adaptations would be needed. The intellectual abilities assessed by the instrument were assumed to be displayed by and salient among the target children. The instrument minimizes the use of language by using simple language forms. Thus, language-driven adaptations were not expected. The limited role of language also minimizes the need for culture-driven adaptations. The other two types of adaptations, theory-driven and familiarity/recognizability-driven adaptation, were more urgent, difficult, and required considerable time and effort, as explained below.

Adapting the Subtests of the KABC-II. Two KABC subtests required a theory-driven adaptation: Number Recall and Atlantis. Four subtests required familiarity/recognizability-driven adaptations: Triangles, Rover, Pattern Reasoning, and Story Completion. Two subtests required both types of adaptation: Word Order and Rebus.

Number Recall measures short-term memory. The child is asked to repeat a series of monosyllabic digits (1 to 9, excluding 7) in the same sequence as presented by the examiner, with series ranging in length from two to nine digits. According to Baddeley's phonological loop model (Baddeley et al. 1975), the number of items that can be stored in short-term memory varies with their phonological length, usually operationalized by the number of syllables. Shorter items are recalled more easily. Thus, Number Recall will be more sensitive to differences in memory capacity when shorter digits are used. As a result, maintaining a constant phonological digit length is important. These theoretical concentrations had a major impact on the instrument in Kannada. All digits in this language are bisyllabic, except 2 and 9 that have three syllables. Therefore, the bisyllabic digits in the Kannada version were used as much as possible. Given their increased difficulty, trisyllabic digits were used only toward the end of the test. As a consequence of this choice, its items differ from the original source test

Atlantis measures long-term storage and retrieval. The child is taught nonsense names for pictures of fish, plants, and shells and then asked to point to the corresponding picture in an array of pictures when it is named. The test requires the child to memorize new phonological information without the support of the meaning or context of the words. Results from a pilot study indicated that the English nonsense names were difficult to distinguish for the Indian children. Therefore, the English nonsense names were replaced with Kannada nonsense names that were sufficiently distinct for easy discrimination by the children.

Triangles assesses visual processing. The child is given several identical foam triangles (blue on one side and yellow on the other) and is asked to put them together so that their final assembled form is identical to a target picture of an abstract design. For easier items, the child assembles a set of colorful plastic shapes to match a model constructed by the examiner or shown in the test booklet. Results from a pilot study indicated that the children found this task very difficult as there are no similar tasks in their daily environment (e.g., jigsaws or other tasks where the spatial orientation of components is important in completing an assembly task). Consequently, the examples and first items on the instrument were simplified to ensure they could be understood and solved by the children. In addition, time constraints were relaxed. These adaptations illustrate the influence that a lack of familiarity with tasks that resemble those on a test can have on test performance.

Rover measures visual processing. The child is asked to move a dog toy (named Rover) to a bone on a checkerboard-like grid that contains obstacles by making as few moves as possible. Pilot study data showed that the children used the direction in which the dog was facing as a cue to how they could or were supposed to start. This finding was both undesirable and unanticipated. Thus, the dog was replaced with a pawn that did not have a directional indication for the children.

Children did not have much experience with similar board games. As a consequence, not all children in the pilot test understood the moves the dog could make on the board. A sample item was changed, and two regular test items were converted to sample items to help ensure that the child understood the principles of the test. As in Triangles, the original time limits were relaxed.

Pattern Reasoning assesses fluid reasoning. Children are shown a series of stimuli that form a logical sequence; however, one stimulus in the series is missing. The child completes the pattern by selecting the correct stimulus from an array of four to six options depicted at the bottom of the page. Most stimuli are abstract, geometric shapes, whereas some easy items use meaningful pictures. The original version requires the assessment of response times at item level. However, results from field tests revealed the difficulty of accurately timing responses that often take only a few seconds. Therefore, scoring was changed to the number of items solved correctly.

Story Completion also measures fluid reasoning. A row of pictures that tell a story is shown but some pictures are missing. The child is given a set of pictures, selects the ones that are needed to complete the story, and places the missing pictures in their correct sequence. This subtest created insurmountable adaptation problems due to the numerous explicit and implicit references to cultural aspects that were unfamiliar or unknown to the children (e.g., having a birthday party, blowing up balloons, and the use of napkins). Therefore, all illustrated pictures were replaced by new items based on those from the Picture Arrangement subset from the Wechsler Intelligence Scale for Children (Wechsler 1949, 1974, 1991). Each Picture Arrangement item consists of a series of pictures that depict a story and are presented in an incorrect order. The child is asked to arrange them in an order that makes a sensible story.

Although Picture Arrangement seemed to be less related to a specific cultural context, the items still needed considerable modification. The number of cards in each item was kept similar to the original Wechsler scales whenever possible. Five new themes were introduced (two sample items and three test items), one item from the original WISC was used, one item from WISC-III was used, and eight items of WISC-R were adapted. The original Picture Arrangement subtest had only one sample item that did not require the child's active engagement. Instead the examiner arranges the cards in the correct order, tells the story, and asks the child whether he or she understood the item. Two sample items that required the child's active engagement were added.

Word Order examines short-term memory. The examiner first names a set of items that the child cannot see and the child then is asked to point to a series of silhouettes of common objects in the same order the examiner named them. On the more difficult tasks, an interference task (i.e., saying the names of colors) is added between the stimulus and the response. Stimuli from the source version of Word Order were selected carefully to ensure that young children with normal receptive language development would readily identify and label all pictures in an adequate manner. The source test only uses objects with one-syllable names to control for phonological length and complexity similar to what was previously observed for Number Recall (theory-driven adaptation). Unlike Number Recall, Word Order does not require the child to respond verbally.

Identifying daily used objects that have one-syllable names in Kannada was difficult. Thus, objects with bisyllabic names were selected (theory-driven adaptation). The additional criteria for selecting new stimuli were that their names and corresponding visual representation (black-and-white drawings) should be unambiguous and highly familiar (familiarity/recognizability-driven adaptation). One out of the twelve original stimuli needed to be redrawn (i.e., the original drawing of a house contained a chimney—a feature unknown to the Indian children). Six of the twelve original stimuli had to be replaced. Drawings of a star, key, hand, moon, heart, and shoe were replaced by drawings of a flower, book, leg, sun, chair, and bus, respectively. The color interference task (color naming) was used to measure recall after injecting an interference task. Children had difficulty naming gray blocks because Kannada does not have a word for gray. Thus, this problem was avoided by using blocks with more familiar colors.

Rebus measures associative memory, verbal learning, and long-term storage and retrieval. The examiner states a word or concept and associates it with a particular drawing. Then the child "reads" aloud phrases and sentences composed of these drawings (e.g., six different drawings can form the sentence "The girl and boy play games"). The research team decided to not use the Rebus because translating and adapting it would not have been possible. The sentences to be produced are related strongly to the specifics of the local language (e.g., the word order in sentences), so that a close (literal) translation was not possible and any modification would produce a considerably different version from the original item. Therefore, the Verbal Learning Test based on the *Rey Auditory Verbal Learning Test* (Rey 1964) replaced the Rebus test. This new test measures immediate memory, efficiency of learning, and recall after short and long delays.

From Test Development to "Testing the Test". As can be seen from the above description, extensive adaptations were needed to design and evaluate a test battery that presumably assessed the intellectual skills measured by the original KABC-II. The adaptations were conducted using iterative procedures. Each iteration included contacts with local psychologists who had expert knowledge in the local language and children's environments. The second component of each iteration involved the administration of subtests to samples of children in the target group. The adaptation process took several months and led to many major and minor changes to the original instrument. The adaptations were made by a team that included persons with cultural, linguistic, and psychometric expertise. The absence of any one type of expertise would have led to a different, presumably less adequate instrument.

The current thinking on test adaptations (e.g., Hambleton and Lee 2013; Hambleton et al. 2005) emphasizes the importance of using mixed methods when engaged in test adaptation. During the design stage, contacts with experts, piloting, and cognitive interviewing are important tools to develop adequate test adaptations. However, these usually qualitative procedures should be complemented by acquiring quantitative data with the new instrument. These data form the core of the quantitative stage that are used to investigate whether and to what extent the psychometric properties of the original instrument are retained in the new instrument.

The adapted KABC-II was administered to 598 children (Malda et al. 2010). Subtest data were reliable, with values ranging from .70 to .96. Validity was examined using confirmatory factor analysis. Findings largely replicated the CHC model underlying the original KABC-II; support for Das-Luria's Planning, Attention-Arousal, Simultaneous, and Successive (PASS) model was not tested. Relations with demographic characteristics (e.g., sex and age) and a measure of arithmetic were consistent with expectations. Thus, the adequacy of the adaptation was supported.

In summary, the adaptation process comprised a small number of large decisions and a large number of small decisions. Large decisions concerned issues about the suitability of the subtests. For example, the need to reject and replace the Rebus occurred early in the adaptation process. However, considerable time was devoted to making numerous refinements to the instrument. For example, multiple iterations with sometimes minute changes were needed to develop an adequate version of Triangles. Similarly, designing, drawing, and refining items of Picture Arrangement required much time before the drawings were clear to all children.

Modifications to the IQ Testing Process

Various other procedures and instruments have been proposed and implemented to increase test validity when testing members of culturally diverse communities. These include testing the limits (e.g., following the standardized administration, the examiner alters procedures to determine if the individual can successfully complete items with suspended time limits), demonstrating and teaching the task to encourage correct performance, incorporating nonverbal measures, and altering and in other ways modifying standardized instructions to obtain qualitative information. Although these alterations jeopardize a test's standardized administration policies, they provide an astute and informed psychologist the opportunity to use the test framework to draw clinical inferences about the person's cognitive abilities. Further modifications to the assessment process as well as test adaptations and alternative methods are described in Table 22.2. Further discussion of "testing the limits" can be found in Sattler (2008). The following provides a brief overview of some test adaptations that may assist in the assessment of intelligence when culture, language, and cultural differences may be salient.

Integration of Modified Instructional Procedures. The WISC-IV Integrated (Kaplan et al. 2004) goes beyond the traditional administration of the standardized format to enable a clinician to gather information regarding underlying processing issues that may impact performance on its 16 subtests. For example, the standard vocabulary subtest of the WISC-IV can be further supplement with both a multiple choice and picture vocabulary versions. For example, a young client who was referred for possible intellectual retardation and tested on the WISC-IV revealed a pattern of scores on the vocabulary subtest that went from 3 to 7 to 9 on the standard, multiple choice and picture vocabulary versions, respectively. Follow-up diagnosis suggested she was an anxious girl with elective mutism and other clinical features that masked her cognitive abilities under standard testing conditions.

Nonverbal Measures. Tests that reduce language or are fully nonverbal often are used with non-English-speaking (or those with limited English proficiency) persons. An increase in the number of non-English-speaking immigrants has increased the need for nonverbal measures of intelligence. Historically, measures of visual conceptual or perceptual organizational ability found in existing tests (e.g., the subtests that previously comprised the Wechsler scale's Performance IQ) were used for this purpose. However, this practice now is deemed inadequate given their reliance upon verbal directions. Nonverbal testing requires an absence of verbal instructions and verbal responses (McCallum and Bracken 2013; Naglieri 2003). The use of nonverbal tests with persons from racial and ethnic minority groups decreases yet does not eliminate IQ differences. Thus, these scales may be referred to as culturally reduced measures.

Cross-Battery Assessment. As noted earlier, cross-battery assessment methods may provide important information regarding the use of intelligence tests with diverse cultural groups through the use of the culture-language interpretive matrix (C-LIM; Flanagan et al. 2013; Ortiz 2013) and culture-language test classifications (C-LTC). This system classifies each subtest within a test battery in light of its cultural loading and linguistic demand (i.e., low, medium, high). For example, the Wechsler scales' Matrix Reasoning subtest is rated low on both linguistic demand and cultural loading. However, the Information, Similarities, Vocabulary, and Comprehension subtests are rated high on both linguistic demand and cultural loading. The classifications are based upon empirical data, when available, or expert consensus procedures and judgment. The use of these systems may assist a clinician when selecting tests and interpreting their data.

Table 22.2 Potential indicators and measures of intelligence for individuals from diverse cultural backgrounds

Options	Instruments/process
Use of "other" markers of intelligence	*Developmental milestones* (e.g., motor and speech development in infancy linked to cognitive functioning; e.g., Murray et al. 2007)
	Achievement indicators (e.g., IQ is correlated: .50 with grades; .55 with total years of education; e.g., Neisser et al. 1996)
Tests of mental chronometry (e.g., reaction time)	People with higher IQ "apprehend, scan, retrieve, and respond to stimuli more quickly than those who score lower" (e.g., Neisser et al. 1996, p. 83)
Curriculum-based assessment (CBM)	CBM may be a viable alternative to traditional norm-referenced testing for African American students (e.g., Fore et al. 2006). CBM provides meaningful information regarding a student's performance with direct linkage to instructional programs
Dynamic assessment	Dynamic assessment and instruction yields findings reducing the minority performance achievement gap in math (e.g., Jeltova et al. 2011)
Use of performance/nonverbal reasoning subtests	Example: use of perceptual reasoning index of Wechsler scales
Use of interpreter	Trained interpreters for the process of assessment may be used with caution given that issues of equivalence (as noted earlier) will impact findings
Use of nonverbal measures	Nonverbal intelligence tests such as the University Nonverbal Intelligence Test (UNIT; 1998); Test of Nonverbal Intelligence Fourth Edition (TONI; Brown et al. 2010); Leiter International Performance Scale Third Edition (Leiter-3; Roid et al. 2013); Raven's Standard Progressive Matrices (RSPM; Raven et al. 1998). Note some nonverbal measures contain verbal instructions while others do not. Nonverbal measures are identified as culturally reduced given absence of language component
Use of measures translated and standardized in country of origin	Many instruments such as the Wechsler Scales have been translated and standardized in many countries (e.g., Georgas et al. 2003)

Biocultural Model. The biocultural model of intelligence (Armour-Thomas and GoPaul-McNicol 1998) incorporates a pre-assessment and ecological evaluation processes to obtain information regarding health, linguistics, and prior experiences in a person's school, home, and community (e.g., educational, psychosocial, acculturation). Attention to these issues may enable a clinician to understand an examinee's personal and environmental contexts. Using this model, psychometric tests first are administered using standardized procedures and then unstandardized methods designed to identify the examinee's intellectual potential are used by suspending the administration time, contextualizing language vocabulary (e.g., allowing the examinee to use the word in a sentence if they are unable to provide a definition), allowing use of paper and pencil, and applying dynamic assessment procedures (test-teach-retest). Thus, in many ways, this process is akin to testing the limits yet goes beyond many of the more routine procedures. Attention to other forms of "intelligences" (e.g., musical) also is included. For example, the work of Howard Gardner reminds us that our views of intelligence sometimes are narrow, in part, because we seemingly judge the test's ability to predict school achievement as the expression of a test's validity. For example, a First Nations student who was failing in school, withdrawn, and hostile toward teachers also displayed athletic skills. Her track and field activities (e.g., short- and long-distance running as well as broad and high jumping) and team sport activities in volleyball and soccer were better than her school peers. An assessment that focused on the use of her skills and interest areas in a learning context found that she could read, communicate, strategize, plan, and keep time charts to determine improvements in performance. She also was able to illustrate human movement and motion using line drawings that were quite exceptional. Thus, the collection and incorporation of additional information gathered through these alternative procedures that may include informal assessment, observation, and interviews may enable a clinician to form a more accurate judgment of an individual's potential and estimated abilities beyond that provided through the use of standardized administrative processes.

Conclusions

The structure of intelligence is universal; there is compelling evidence through many cross-cultural studies of intelligence tests that these measures are both robust and invariant. Basic features (e.g., working memory, spatial reasoning) are part of the cognitive repertoire of individuals across the globe. However, this universality could be misconstrued if applied universally in a rigid manner. While we may agree that crystallized intelligence reflects the capacity to learn, store, and retrieve information (the terms *schema, assimilation, accommodation*, declarative, *semantic* and *procedural knowledge* could be used), the examples in this chapter show that beyond the underlying processes and structure of intelligence that is manifested and what is valued will vary across cultures. Thus, persons engaged in the assessment of intelligence cross-nationally should consider the following four issue: the nature and importance of intellectual abilities desired by a culture or society; whether these qualities differ based on a person's age, gender, and socioeconomic status; the degree of overlap between these desired intellectual qualities and the current structure of intelligence as more universally defined such as the CHC model; and the availability of valid methods to assess intelligence in light of this information.

The repertoire of abilities captured by the structure of intelligence is not used uniformly by all people. One's culture and socialization practices contribute to the promotion of specialized and possibly different abilities. Persons acquire culture-specific skills along with universal skills through their socialization. Cultures also reinforce the display of biologically based abilities. As a result, assessment procedures developed in one cultural context may not transfer well to another context.

Knowledge of the assessment procedures used as immigrants were entering the United States at Ellis Island may help avoid similar practices (Gould 1996; Tulsky et al. 2003). The inadequacy of the testing procedure combined with the exhaustion of the immigrants after the long boat trip, the effects of fleeing from war and poverty, the unfamiliarity of this new context, and the stress of wanting an opportunity for a better life underscores the importance of assessing intellectual skills in ways that acknowledge the examinee's social, cultural, and linguistic background and align assessment methods consistent with it along with the context in which the data will be used. Many intelligence tests have been developed for Western cultural groups and thus may be culturally loaded. A test's good psychometric properties do not compensate for their use in culturally insensitive ways. A theory is neither right nor wrong, neither good nor bad, and should be judged whether it is useful or useless in understanding, predicting, and even changing human behavior. The tools we create to reflect the latent traits and human characteristics derived from these theories and models must be viewed in the same light. While one could argue with this viewpoint (e.g., Nazi views of Aryan racial superiority leading to support for ethnic cleansing and genocide programs or restricting opportunity based on racial views of intelligence), it seems much more meaningful if also placed in the context of ethical practice. Psychologists who use measures of intelligence tests have an important responsibility to the individual client and to society, one that should reflect comprehensive and relevant training and in-depth knowledge of psychological research as well as practice standards, cultural sensitivity, and ethics. To paraphrase Anne Anastasi whose brilliant writings on psychological testing and assessment informed generations of psychologists and others, "we are all of equal value as human beings" (Gottfredson 1997). And if we superimpose on that, the recognition echoed by Kluckhohn et al. (1953) that noted the ways in which we are "like all others, some others and no others," the construct of intelligence and its assessment in a culturally diverse world will be well served. The information in this chapter is intended to identify and clarify issues that may arise when using measures of intelligence cross-culturally or cross-nationally.

Appendix A

International Test Commission Guidelines

The use of cognitive and other ability measures can engender strong negative feelings and opinions—particularly in reference to their misuse with groups that may display lower than average mean scores, when tests are misused when assessing immigrants and second-language learners, or when tests do not meet suitable psychometric standards. During the 1970s and 1980s, articles in newspapers and professional journals alleged that examiners often are poorly prepared; tests typically rely on a small sample of behavior, often those to which persons from lower-class homes lacked exposure; tests overlook functional daily used qualities; tests are the servants of the established middle class and are designed to discriminate against the poor and colored; and that cheating on tests is widespread and invalidates their use. In short, the allegations suggest tests generally are unfair, culture bound, biased, and used discriminatorily (Oakland 1977; Phelps 2008).

Many of these issues surfaced in countries other than the United States during this same period (Oakland 2009; Oakland et al. 2013) fueled, in part, by two conditions: (1) the use of pirated copies of tests, developed originally in the United States or Western Europe, and translated into the local language and without adequate efforts to acquire norms or validity data as well as (2) strong reliance on theory from social psychology that expressed disdain for the principle of meritocracy and favoritism toward egalitarian views.

The International Test Commission (ITC) was established during this time to promote an exchange of information on test development and use as well as to address prevailing issues that cut across national boundaries and cultures (Oakland et al 2001). Thus, given its international focus, ITC leadership recognized its need to attend to prevailing issues that were impeding test development and use regionally and internationally. The lack of authoritative guidelines can contribute to assessment services that are unreliable, inferior in quality, and reflect unsuitable standards. Thus, the ITC began developing various guidelines intended to promote sound testing practices. Features of four guidelines that may be most pertinent to cognitive assessment are summarized below. Information on the guidelines can be found on the ITC website at http://www.intestcom.org.

Test Adaptation Guidelines. The ITC recognized most countries lack resources needed to develop their own tests and thus will continue the practice of obtaining tests, often measures of cognitive ability (Hu and Oakland 1991), from others. Thus, guidelines were needed to adapt these tests for local use. The guidelines provide guidance in reference to the context of a test's use, test development and adaptation, test administration, and documentation/scoring interpretations. These guidelines distinguish two processes: translating/adapting existing tests and instruments as well as developing new instruments intended to be used internationally to provide international comparisons.

ITC Guidelines for Test Use. These guidelines were developed to help overcome unsuitable assessment practices by focusing on test user competencies. Competencies require acquiring basic knowledge (e.g., psychometric principles and procedures along with technical requirements of tests) and skills (e.g., suitable use of assessment procedures), displaying suitable professional and ethical behaviors, assuming responsibility for test use, ensuring test materials are secured, ensuring that test results are treated confidentially, evaluating the potential utility of testing, selecting technically sound tests appropriate for the situation, considering issues of fairness, establishing rapport, scoring and analyzing test results accurately displaying a good understanding of the test's theoretical or conceptual basis, communicating the results clearly and accurately to relevant others, and reviewing the appropriateness of the test and its use.

International Guidelines on Computer-Based and Internet-Delivered Testing. The use of computers to administer scores and report tests has become somewhat routine. Lamentably, individuals with limited ability to develop suitable tests are able to market their inferior products online and potentially reach all people who have Internet access (Friedman 2005). These guidelines are intended to promote good practices for computer-based (CBT) and Internet-delivered testing. They address technical issues (e.g., hardware and software requirements; the robustness of the CBT/Internet test; human factors in the presentation of materials via a computer or the Internet; providing reasonable adjustments to the test's technical features for test-takers who display disabilities; and providing help, information, and practice items) as well as quality control issues (e.g., knowledge, competence, and appropriate use of CBT/Internet testing; the psychometric qualities of CBT/Internet tests; equivalence between paper-and-pencil tests and those delivered via CBT/Internet methods; scoring and analyzing CBT/Internet results accurately; interpret results appropriately, including feedback; and equality of access for all groups).

Proposed ITC Guidelines for the Clinical Assessment of Immigrants and Second-Language Learners. An increase in personal mobility has led to an increase in the number of persons who hold immigrant status, especially in Western countries. They frequently seek improved lifestyles, including education, work, safety, and other conditions that contribute to their health, welfare, and stability. Their immigrant status implies they are residing in a different culture and are likely to lack fluency in the host country's primary or preferred language and thus are second-language learners. Attempts to assess their cognitive abilities pose severe challenges to assessment specialists. These proposed guidelines consider five issues that may impact the clinical assessment of persons whose cultural and linguistic qualities differ from the local culture: qualities associated with culture and language, four personal qualities that commonly are assessed (i.e., medical, social, cognitive, and behavioral), the importance of psychometric qualities, interpersonal relationships, and ethical issues. The importance of considering the acculturation status of immigrant clients (i.e., their orientation and level of adjustment to their ethnic culture and the mainstream culture) and possible ways to address acculturation issues are described. Possible test modifications together with best practice guidelines are suggested. These proposed guidelines are intended to assist professionals in the provision of clinical assessment services for persons who are acculturating to a new culture and are second-language learners. The ITC leadership currently is reviewing these proposed guidelines.

References

Armour-Thomas, E., & GoPaul-McNicol, S. A. (1998). Assessing intelligence: Applying a biocultural model. Thousand Oaks, CA: Sage.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575–589.

Berry, J., Poortinga, Y., Breugelmans, S., Chasiotis, A., & Sam, D. (2011). *Cross-cultural psychology. Research and applications* (3rd ed.). Cambridge, UK: Cambridge University Press.

Bracken, B. A., & McCallum, R. S. (1998). *Universal nonverbal intelligence test examiner's manual*. Itsaka: Riverside.

Brislin, R. W. (1970). Back translation for cross-cultural research. Journal of Cross Cultural Psychology, 185–216.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2010). *Test of nonverbal intelligence* (4th ed.). San Antonio: Pearson.

Butcher, J. N., & Han, K. (1996). Methods of establishing cross- cultural equivalence. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 44–63). Minneapolis: University of Minnesota Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1940). A culture free intelligence test, I. *Journal of Educational Psychology, 31*, 176–199.

Dasen, P. R. (2011). Culture, cognition and learning. In A. Bame Nsamenang & T. S. Tschombe (Eds.), *Handbook of African educational theories and practices: A generative teacher education curriculum contents* (pp. 187–202). Bamenda: Human Development Resource Centre.

Edwards, O., & Oakland, T. (2006). Factorial invariance of Woodcock-Johnson scores for African Americans

and Caucasian American. *Journal of Psychoeducational Assessment, 24*, 358–366.
Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (2nd ed.). Hoboken: Wiley.
Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century*. New York: Cambridge University Press.
Fore, C., Burke, M. D., & Martin, C. (2006). Curriculum-based measurement: An emerging alternative to traditional assessment for African American children and youth. *The Journal of Negro Education, 75*, 16–24.
Friedman, T. (2005). *The world is flat*. New York: Farrar, Straus and Giroux.
Georgas, J., Weiss, L., Van de Vijver, F. J. R., & Saklofske, D. H. (Eds.). (2003). *Cultures and children's intelligence: A cross-cultural analysis of the WISC-III*. New York: Academic.
Goodnow, J.1. (1984). On being judged intelligent. International Journal of Psychology, 19:391–406.
Gottfredson, L. S. (1994). The science and politics of race-norming. *American Psychologist, 49*, 955–963.
Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence, 24*(1), 13–23.
Gould, S. J. (1996). *The mismeasure of man*. New York: Norton & Company.
Grigorenko, E. L., Geissler, P. W., Prince, R., Okatcha, F., Nokes, C., Kenny, D. A., Bundy, D. A., & Sternberg, R. J. (2001). The organization of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development, 25*, 367–378.
Hambleton, R. K., & Lee, M. K. (2013). Methods for translating and adapting tests to increase cross-language validity. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The Oxford handbook of child psychological assessment* (pp. 172–181). New York: Oxford University Press.
Hambleton, R., Merenda, P., & Spielberger, C. (Eds.). (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah: Erlbaum.
Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing. *American Psychologist, 47*, 1083–1101.
Helms-Lorenz, M., Van de Vijver, F. J. R., & Poortinga, Y. H. (2003). Cross-cultural differences in cognitive performance and Spearman's Hypothesis: G or C? *Intelligence, 31*, 9–29.
Hernstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.
Hess, R., Kashiwagi, K., Azuma, H., Price, G., & Dickson, W. P. (1980). Maternal expectations for mastery of developmental tasks in Japan and the United States. International Journal of Psychology, 15, 259–271.
Hu, S., & Oakland, T. (1991). Global and regional perspectives on testing children and youth: An international survey. *International Journal of Psychology, 26*, 329–344.
International Test Commission, (2010). International test commission guidelines for translating and adapting tests. Retrieved from http://www.intestcom.org/upload/sitefiles/40.pdf
Jeltova, I., Birney, D., Fredine, N., Jarvin, L., Sternberg, R. J., & Grigorenko, E. L. (2011). Making instruction and assessment responsive to diverse students' progress: Group administered dynamic assessment in teaching mathematics. *Journal of Learning Disabilities, 44*, 381–395.
Jensen, A. (1972). *Genetics and education*. San Francisco: Harper & Row.
Jensen, A. (1973). *Educability & group differences*. San Francisco: Harper & Row.
Jensen, A. (1980). *Bias in mental testing*. New York: The Free Press.
Kaplan, E., Fein, D., Kramer, J., Delis, D., & Morris, R. (2004). *WISC-IV integrated*. San Antonio: The Psychological Corporation.
Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children, second edition: Manual*. Circle Pines: AGS Publishing.
Kluckhohn, C. E., Murray, H. A., & Schneider, D. M. (Eds.). (1953). *Personality in nature, society, and culture*. New York: Knopf.
Lonner, W. J. (1980). The search for psychological universals. In H. C. Triandis & W. W. Lambert (Eds.), *Handbook of cross-cultural psychology* (Vol. 1, pp. 143–204). Boston: Allyn and Bacon.
Lonner, W. J. (1985). Issues in testing and assessment in cross-cultural counseling. *Counseling Psychologist, 13*, 599–614.
Lonner, W. J. (2011). The continuing challenge of discovering psychological 'order' across cultures. In F. J. R. Van de Vijver, A. Chasiotis, & S. M. Breugelmans (Eds.), *Fundamental questions of cross-cultural psychology* (pp. 64–94). Cambridge, UK: Cambridge University Press.
Malda, M., Van de Vijver, F. J. R., Srinivasan, K., Transler, C., Sukumar, P., & Rao, K. (2008). Adapting a cognitive test for a different culture: An illustration of qualitative procedures. *Psychology Science Quarterly, 50*, 451–468.
Malda, M., Van de Vijver, F. J. R., Srinivasan, K., Transler, C., & Sukumar, P. (2010). Traveling with cognitive tests: Testing the validity of a KABC-II adaptation in India. *Assessment, 17*, 107–115.
Marfo, K. (2011). Envisioning an African child development field. *Child Development Perspectives, 5*(2), 140–147.
Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic mix of traits? *American Psychologist, 63*, 503–517.
McCallum, R. S., & Bracken, B. A. (2013). Intelligent testing of underserved populations. In J. C. Kaufman (Ed.), *Intelligent testing: Integrating psychological theory and clinical practice* (pp. 193–219). New York: Cambridge University Press.
McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P.

Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136–181). New York: The Guilford Press.
McGrew, K., & Flanagan, D. (1998). *The intelligence test desk reference: Gf-Gc cross-battery assessment*. Boston: Allyn & Bacon.
Mercer, J. R., & Lewis, J. F. (1977). *SOMPA, System of multicultural pluralistic assessment: Parent interview manual*. New York: Psychological Corporation.
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.
Mpofu, E. (2004). Being intelligent with Zimbabweans: A historical and contemporary view. In R. J. Sternberg (Ed.), *International handbook of intelligence* (pp. 364–390). Cambridge/New York: Cambridge University Press.
Mpofu, E., & Nyanungo, K. R. (1998). Educational and psychological testing in Zimbabwean schools: Past, present and future. *European Journal of Psychological Assessment, 14*, 71–90.
Mpofu, E., & Ortiz, J. (2009). Equitable assessment practices in diverse contexts. In E. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 41–76). New York: Springer.
Mpofu, E., Gudyanga, E., & Ngara, C. (2007). Constructions of giftedness among The Shona of Central-Southern Africa. In S. N. Phillipson & M. McCann (Eds.), *What does it mean to be gifted? Sociocultural perspectives* (pp. 225–252). Mahwah: Lawrence Erlbaum.
Mpofu, E., Ntinda, K., & Oakland, T. (2012). Understanding human abilities in Sub-Saharan African settings. *Online Readings in Psychology and Culture*, Unit 4. Retrieved from http://scholarworks.gvsu.edu/orpc/vol4/iss3/2. ISBN 978-0-9845627-0-1
Mpofu, E., Oakland, T., Ntinda, K., Seeco, E., Maree, J.G. (2014). Constructing a framework for the use of tests within a developing country's school system. International Perspectives in Psychology: Research, Practice, Consultation, 3(2), 106–122. doi: 10.1037/ipp0000015
Murphy, L. M., Geisinger, K. F., Carlson, J. F., & Spies, R. A. (Eds.). (2011). *Tests in print VIII*. Lincoln: Buros Institute of Mental Measurements.
Murray, G. K., Jones, P. B., Kuh, D., & Richards, M. (2007). Infant developmental milestones and subsequent cognitive function. *Annals of Neurology, 62*(2), 128–136.
Naglieri, J. A. (2003). *Naglieri nonverbal abtlity Test*. San Antonio: Psychological Corporation.
Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77–101.
Nisbett, R. E. (2009). *Intelligence and how to get it: Why schools and cultures count*. New York: Norton.
Nsamenang, A. B. (2006). Human ontogenesis: An indigenous African view on development and intelligence. *International Journal of Psychology, 41*, 293–297.
Oakland, T. (Ed.). (1977). *Psychological and educational assessment of minority children*. Larchmont: Brunner-Mazel.
Oakland, T. (2009). How universal are test development and use? In E. Grigorenko (Ed.), *Assessment of abilities and competencies in an era of globalization* (pp. 1–40). New York: Springer.
Oakland, T., & Harrrison, P. (2008). Adaptive behavior and skills: An introduction. In T. Oakland & P. Harrison (Eds.), *ABAS-II: Clinical uses and interpretations*. New York: Elsevier.
Oakland, T., & Hu, S. (1992). The top ten tests used with children and youth worldwide. *Bulletin of the International Test Commission, 19*, 99–120.
Oakland, T., Poortinga, Y., Schlegel, J., & Hambleton, R. K. (2001). International Test Commission: Its history, current status, and future directions. *International Journal of Testing, 1*, 3–32.
Oakland, T., Wechsler, S., & Maree, K. (2013). Test use with children across cultures: A view from three countries. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology* (Vol. 3, pp. 231–258). Washington, DC: American Psychological Association.
Ortiz, S. O. (2013). *Distinguishing difference from disorder in English learners and guidance on how to use the Culture-Language Interpretive Matrix*. New York: Cross-Battery Institute, Fordham University.
Pence, A. (2011). Early childhood care and development research in Africa: Historical, conceptual, and structural challenges. *Child Development Perspectives, 5*(2), 112–118.
Phelps, R. (Ed.). (2008). *Correcting fallacies about educational and psychological testing*. Washington, DC: American Psychological Association.
Porteus, S. D. (1937). Primitive intelligence and environment. New York: MacMillan Co.
Prifitera, A., Saklofske, D. H., & Weiss, L. G. (Eds.). (2008). *WISC-IV clinical interpretation and intervention*. San Diego: Academic.
Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven's standard progressive matrices – Parallel form*. Oxford: Oxford Psychologists Press.
Rey, A. (1964). *L'examen clinique en psychologie [The clinical examination in psychology]*. Paris: Presses Universitaires de France.
Roid, G. H., Miller, L. J., Pomplum, M., & Koch, C. (2013). *Leiter international performance scale* (3rd ed.). Torrance: Western Psychological Services.
Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*(2), 235–294.
Ryan, C. (2013, August). *Language use in the United States 2011: American community survey reports.* Retrieved from www.census.gov/prod/2013pubs/acs-22.pdf

Saklofske, D. H., Reynolds, C. R., & Schwean, V. L. (Eds.). (2013). *Oxford handbook of child psychological assessment*. New York: Oxford University Press.
Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego: Jerome Sattler.
Serpell, R. (2008). Participatory appropriation and the cultivation of nurturance: A case study of African primary health science curriculum development. In P. R. Dasen & A. Akkari (Eds.), *Educational theories and practices from the "majority world"* (pp. 71–97). New Delhi: Sage.
Serpell, R. (2011). Social responsibility as a dimension of intelligence, and as an educational goal: Insights from programmatic research in an African society. *Child Development Perspectives, 5*, 126–133.
Serpell, R., & Haynes, B. P. (2004). The cultural practice of intelligence testing: Problems of international export. In R. J. Sternberg & E. L. Grigorenko (Eds.), *Culture and competence: Contexts of life success* (pp. 163–185). Washington, DC: American Psychological Association.
Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media and public policy*. New Brunswick: Transaction Publishers.
Sternberg, R. J. (1985). Implicit theories of intelligence, creativity and wisdom. *Journal of Personality and Social Psychology, 49*, 607–627.
Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology, 3*, 292–375.
Sternberg, R. J., & Jordan, J. (Eds.). (2005). *A handbook of wisdom: Psychological perspectives*. New York: Cambridge University Press.
Sternberg, R. J., Nokes, C., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence, 29*, 401–418.
Trimble, J. E. (2010). Cultural measurement equivalence. In C. C. Clauss-Ehlers (Ed.), *Encyclopedia of cross cultural school psychology* (pp. 316–318). New York: Springer.
Tulsky, D., Saklofske, D. H., Chelune, G. J., Heaton, R. K., Ivnik, R. J., Bornstein, R., Prifitera, A., & Ledbetter, M. F. (2003). *Clinical interpretation of the WAIS-III and WMS-III*. San Diego: Academic.
Van de Vijver, F. J. R. (1997). Meta-analysis of cross-cultural comparisons of cognitive test performance. *Journal of Cross-Cultural Psychology, 28*, 678–709.
Van de Vijver, F. J. R., & Leung, K. (1997). Methods and data analysis for cross-cultural research. Newbury Park, CA: Sage.
Van de Vijver, F. J. R. (2001). The evolution of cross-cultural research methods. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 77–97). Oxford: Oxford University Press.
Van de Vijver, F. J. R. (2003). Bias and equivalence: Cross-cultural perspectives. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 143–155). New York: Wiley.
Van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). New York: Cambridge University Press.
Van de Vijver, F. J. R., & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational tests and psychological tests for cross-cultural assessment* (pp. 39–64). Mawhaw: Erlbaum.
VandenBos, G. (2007). *APA dictionary of psychology*. Washington, DC: American Psychological Association.
Wechsler, D. (1949). *Wechsler intelligence scale for children*. New York: Psychological Corporation.
Wechsler, D. (1974). *Manual for the Wechsler intelligence scale for children-revised*. San Antonio: The Psychological Corporation.
Weiss, L. G., Saklofske, D. H., Coalson, D., & Raiford, S. E. (Eds.). (2010). *WAIS-IV clinical use and interpretation*. San Diego: Academic.
Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013a). WAIS-IV and clinical validation of the four and five factor interpretative approaches. *Journal of Psychoeducational Assessment, 31*, 94–113.
Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013b). WISC-IV and clinical validation of the four and five factor interpretative approaches. *Journal of Psychoeducational Assessment, 31*, 14–131.
Wober, M. (1974). Towards an understanding of the Kiganda concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 261–280). London: Methuen.
Wolf, T. (1973). *Alfred Binet*. Chicago: University of Chicago Press.
Yan, G., & Saklofske, D. H., (2004). Intelligence: Views of Chinese psychologists. Paper presented at 28th international congress of psychology, August. Beijing, China.
Yan, G., Saklofske, D. H., & Oakland, T. (2009). Contrasting views of intelligence: Chinese and North American perspectives. *School Psychology International, 30*, 456–473.

23 Common and Variable Aspects of Intelligence

Arthur MacNeill Horton Jr. and Cecil R. Reynolds

Intelligence has been conceptualized as a product of the overall physiological efficiency of the brain itself and crucial for adaptive problem solving. David Wechsler (1944) defined intelligence as "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment" (p. 3). This chapter considers the assessment of human intelligence from a broad perspective of brain-behavior relationships as an aid to understanding implications of normal as well as abnormal intellectual ability through the discussion of biobehavioral paradigms related to human intelligence. All psychologists have a strong interest in intelligence from theoretical, social, and clinical perspectives. Level of intelligence is important to establish in multiple contexts as a selection and placement criterion as well as an aid in diagnosis and treatment, including as a baseline of overall mental function against which more specific cognitive skills may be compared. Human intelligence as a clinical biobehavioral concept was initially proposed by Alfred Binet. Essentially, City of Paris, France, public school officials were concerned that children of impaired cognitive ability were not being discriminated from children of normal cognitive ability and that was disruptive to the education of all children because of the need for differential education methods for both groups of children. The education officials requested Binet to develop a method to discriminate children of impaired cognitive ability from children of normal cognitive ability to determine which children should be in special education, and he subsequently created the first successful standardized intelligence test (Binet and Simon 1905, 1908). It is noteworthy that the task was one of practical significance – selection and appropriate placement of children in an academic context, similar to what is done today through the SAT Reasoning Test (previously the Scholastic Aptitude Test and Scholastic Assessment Test) or Graduate Record Exams (GRE).

Later, David Wechsler developed an intelligence test for adults to aid in clinical assessment (Wechsler 1939). Wechsler's older brother (Israel Wechsler) was a neurologist (Chief of Neurology at Bellevue Hospital in New York City), and the

---

Note: This chapter is based in part on prior work of Drs. Reynolds and Horton in Reynolds, C. R., C. Castillo and Horton, A. M. Jr. (2003). The Neuropsychological Basis of Intelligence: Forensic Implications. In A. Horton & L. Hartlage (Eds.), *Handbook of Forensic Neuropsychology*. Second Edition New York: Springer Publishing Company.

Disclosure Statement: Dr. Reynolds is senior author of the Reynolds Intellectual Assessment Scale (RIAS) and the Test of Verbal Conceptualization and Fluency (TVCF). Dr. Horton is second author of the TVCF.

A.M. Horton Jr., Ed.D., ABPP, ABN (✉)
Neuropsychology section, Psych Associates of Maryland, LLC, 120 Sister Pierre Drive, Suite 403, Towson, MD 21204, USA
e-mail: drmachorton@hotmail.com

C.R. Reynolds, Ph.D.
Texas A&M University, 40 N 1H 35, 12B1, Austin, TX 78701, USA
e-mail: crrh@earthlink.net

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_23,

need for an intelligence test for standardized examination for neurological patients may have been an influence. Wechsler had been a research assistant to Robert Yerkes when Yerkes was in charge of assessing the 17 million men who were drafted to fight in World War I. The 17 million men were assessed with psychological tests commonly used at that time to select which men should be selected for officer training (Wechsler 1939). Two tests were developed, the Army Alpha and the Army Beta. Army Alpha was a verbal test of mental abilities used to assess recruits that were native English speakers. After using Army Alpha, it was noted that many recruits were immigrants from Eastern Europe and Italy who were not native English speakers but might be able to be good officers. To assess these nonnative English speakers properly, Army Beta, a nonverbal mental abilities test, was developed. The psychological tests that had been most successful for selecting officers during World War I were used by Wechsler when he developed his own first intelligence test (Wechsler 1939).

Wechsler combined tasks similar to those on the Army Alpha and Beta tests (verbal and nonverbal tests) to form the Wechsler-Bellevue Intelligence Scale (Wechsler 1939). The most important contribution of Wechsler (1939) was methodological as he standardized the administration and scoring of the tests and normed the intelligence test in large national samples of adults (Wechsler 1952) and introduced age-corrected deviation scaled scores for interpretation of intelligence test performance. Clinical psychology as a field has thus had a long-term interest in evaluating intelligence (Horton and Wedding 1984).

The Role of "*g*" In Neuropsychological Models of Intelligence

The Greek philosopher Aristotle first suggested that intellect could be assessed by a single mental ability variable, *nous* (Detterman 1982). As Aristotle was also the tutor of Alexander the Great, Aristotle clearly had some academic experience with an individual of great intelligence – his own student who became the ruler of most of the then known world. The concept of *g* has been conceptualized as the average of an individual's higher-level cognitive abilities as assessed by many different types of cognitive tasks. Put another way, *g* can be conceptualized as a latent trait rather than an observable outcome.

Research on *g* has been extremely important in psychology, and *g* has been a very useful means of conceptualizing overall intellectual ability (Aluja-Fabregat et al. 2000; Kane 2000). Jensen (1998) after reviewing the empirical research for the presence of a general cognitive ability factor in intelligence concluded that if a very large number of tests were used to assess a very wide spectrum of mental abilities, then a *g* factor would always be found (Jensen 1998). Failures to find a *g* factor in prior research studies were attributed to failures to use a large enough number of tests and assess a large enough different types of abilities (Jensen 1998).

Moreover, researchers (Reynolds and French 2003) have suggested the study of *g* and the study of cognitive processing styles are complementary areas in intelligence research investigation. Indeed, evidence for simultaneous and successive information processes in the human brain may be complementary to the concept of a *g* factor. The verbal and performance factors found in research studies of intelligence testing are examples of the related cognitive factors in diverse populations (Reynolds 1981). The abstract concept of *g* can be seen as possibly complementary to differences in the level or efficiency of information processing (Das et al. 1979; Detterman 1982). For example, Travers (1977) and Luborsky et al. (1971), in studies of psychotherapy outcome research, found that the best predictor of successful psychotherapy outcome was the intelligence level (i.e., *g*) of the individual receiving psychotherapy. The researchers unfortunately seem to have overlooked the option of assessing the contribution of variable of the intelligence level of the person delivering the psychotherapy to psychotherapy outcome but admittedly that would be a difficult study to conduct for multiple reasons. In addition, rehabilitative success of brain-injured neurological patients is best predicted by the pre-morbid intelligence level of the individual patients (Golden 1978).

It is noteworthy that the abstract concept of *g* has been considered limited by the biological integrity and physiological efficiency of the human brain (Brand 1996; Vernon 1998). Harmony (1997) and Languis and Miller (1992) have suggested that physiological measures, such as the EEG and/or auditory evoked potentials, could be utilized to assess aspects of cognitive ability. Jensen's research (1978, 1998) on reaction times and evoked potentials had suggested that *g* could be conceptualized as general physiological efficiency of the central nervous system. Future understanding of the concept of *g* will require elucidation of the method and components of information processing in the brain. It is possible that further elucidation of brain-behavior models may contribute to understanding of *g*. Consideration of the relationship of *g* to contemporary models of brain-behavior may prove very helpful.

Luria's Brain-Behavior Model

Alexander R. Luria, a Soviet neurologist and neuropsychologist, had important insights into brain functioning (Horton 1987). Perhaps the most important insight was the concept of the complex functional system (i.e., multiple diverse brain areas subserve particular behavioral abilities) (Horton 1987). Using the cultural-historical theory of brain-behavior relationship, Luria was able to perform an evaluation of an individual's neurological status (Horton 1987). Grossly oversimplified, it could be averred that he (Luria 1973) described sensory and motor functions of the brain as having highly specific functional localizations, while higher-level mental processes required coordination of multiple areas of the brain. In other words, lower-level functions were hardwired in specific neuroanatomical areas but higher-level functions were widely distributed throughout the human brain. Put still another way, higher-level human brain functions require multiple areas of the brain to accomplish complex behaviors, but lower-level human brain functions (i.e., sensorimotor functions) are localized in a specific area of the brain (Reynolds 1981). Higher-level human brain functions are process specific, and processing of information requires coordination of diverse neuroanatomical brain sections (Ashman and Das 1980). Higher-level human brain organization (Luria 1973) further was characterized as the brain's higher-level processing being organized into three major human brain areas. The first human brain area included the brainstem and reticular formation, the midbrain, pons, and medulla. The second human brain area included the parietal, occipital, and temporal lobes (Luria 1973). The third human brain area included all of the cerebral cortex anterior to the sensory-motor strip (i.e., Rolandic fissure). The three major higher-level brain areas (Luria 1973) all function in a dynamic reciprocal interaction to subserve higher-level cognitive processing, or in other words the higher-level processing depends on multiple diverse areas of the human brain (Reynolds 1981). As earlier noted, lower-level functions are more hardwired to specific neuroanatomical brain areas.

The notion of the brain as a dynamic functional system, it should be acknowledged, was first proposed by Hughlings Jackson, an English physician who lived in the nineteenth century, and was further elucidated by Luria (Horton 1987). Essentially, higher mental processes are seen as based on multiple diverse human brain areas communicating and working together, and as a result higher-level functions may be disrupted by the destruction of a communication channel of the functional system (Luria 1964). Further, disturbances of higher-level mental functions can be influenced based on the specific localization of the brain damage (Luria 1964). Therefore, rehabilitation of the human brain higher-level functional system, if there is specific brain damage, will require the brain to assemble an alternative sequence of human brain areas working together to perform specific behavioral tasks in a new way (Luria (1964). The localizing brain area responsible for behavioral disturbance can be determined by qualitatively analyzing the difficulty experienced in performing a specific behavioral task (Luria 1964).

Neuroanatomical Area One

The first neuroanatomical area of the human brain, the brain stem, subserves maintaining consistent arousal, attention, and concentration abilities. The energy level and tone of the entire human cerebral cortex allow a stable platform to organize the various higher-level cognitive functions of the human brain. The brain stem, first neuroanatomical area, includes the reticular formation, the posterior hypothalamic and brainstem portions of the brain. Damage to the first neuroanatomical area of the human brain can cause lowering of the level of consciousness in the human cerebral cortex, disrupting higher-level cognitive functioning thereby giving rise to disorganized behavior.

Neuroanatomical Area Two

The area posterior to the central sulcus (i.e., parietal, occipital, and temporal lobes) is included in the second neuroanatomical area of the brain. The second neuroanatomical area of the human brain is primarily receptive in nature, integrating diverse sensory inputs, storing, integrating, and organizing sensory information. The second neuroanatomical area of the brain allows perception, analysis, and synthesis of sensory stimuli (e.g., auditory in the temporal lobes, visual in the occipital lobes, and tactile in the parietal lobes). Within the second neuroanatomical area of the brain, each lobe sensory stimuli processing (auditory in the temporal lobes, visual in the occipital lobes, and tactile in the parietal lobes) is organized into three hierarchical zones. The *primary zone* perceives and retains incoming sensory stimuli. The *secondary zone* analyzes and organizes sensory information from the *primary zone*. The *tertiary zone* receives sensory information (auditory in the temporal lobes, visual in the occipital lobes, and tactile in the parietal lobes) from the multiple *secondary zones of the three lobes* and organizes the information into higher-level cognitive processes subserving complex human behavior.

Neuroanatomical Area Three

The frontal lobes which involves the initiation, development, and monitoring of plans for behavior are included in the third neuroanatomical area of the human brain. In other words, frontal lobes, the third neuroanatomical area, receive and evaluate organized sensory input from the first and second neuroanatomical areas of the human brain and perform executive functions integrating the information to subserve complex adaptive problem solving in a managerial role (Luria 1973; Obrzut and Obrzut 1982). The frontal lobes, in addition to direct connections to the second neuroanatomical areas, are also directly connected to the reticular formation in the first neuroanatomical area of the brain. This series of reciprocal communication neural networks mediates the activation and processing of higher-level cognitive processing throughout the human cerebral cortex. Performing an executive function, the frontal lobes direct attention and concentration processes in the human brain. The direct connections among the first, second, and third neuroanatomical areas of the human brain facilitate reciprocal neural network communication systems that facilitate complex human decision making and adaptive problem solving based on arousal, attention, and organized sensory input. The coordination of first neuroanatomical area of the brain with the second and third neuroanatomical areas thereby facilitates the initiation, development, and monitoring of behavioral plans and their timely, efficient, and effective evaluation. In contrast with arousal role of the first neuroanatomical area and the receptive role of the second neuroanatomical area, the third neuroanatomical area of the brain has an expressive, generative role. In a nutshell, it is noted that human higher-level cognitive functioning is facilitated by the dynamic and reciprocal interplay of the three neuroanatomical area of the brain (Luria 1964; Golden et al. 1979; Golden 1987; Horton 1987).

Simultaneous and Successive Cognitive Processes

Further elucidation of the functioning of the second neuroanatomical area (Luria 1964) involves appreciation of modes of information processing. These can be characterized usefully as simultaneous and successive (or sequential) cognitive processes. Put another way, sensory information can be processed in sequence or one element at a time in order or simultaneously where all of the information is processed as a whole or as a gestalt, in other words, describing a group of trees as oak, pine, birch, etc., or as a forest. Simply put, sensory stimuli, in the second neuroanatomical area, can be processed through either simultaneous or successive means (Kaufman 1979b). Simultaneous and successive processes can be used to process any specific sensory modality (i.e., auditory, visual, tactual, etc.) or stimulus elements (verbal, nonverbal) (Ashman and Das 1980). Which type of processing, either simultaneous or successive cognitive processing, is most efficiently effective will depend on the task demands, attention demands required by the task, and preferred means for completing the task (Hall et al. 1988; Watters and English 1995; Willis 1985). Verbal communications may be processed effectively through linear successive methods such as dictating or writing a letter. Spatial tasks, such as map reading, may be processed effectively through simultaneous-processing strategies. Or to use another example, a forest ranger might know each type of tree in a forest, but a hiker would be more concerned with the concept of forest.

Simultaneous Processing

This is the synthesis of separate elements into spatially related groups with direct access to any separate element (Das et al. 1979). Within the second neuroanatomical area, the right occipital and parietal lobes of the human brain subserve simultaneous information processing (Naglieri et al. 1983; Willis 1985). Commonly considered measures of simultaneous processing can include visual-spatial ability tests (Kirby and Das 1977).

Successive Processing

In contrast, successive (or sequential) processing is linear accessing of information in a serial fashion (Das et al. 1979). In the second neuroanatomical area, the left temporal lobe of the brain subserves successive (or sequential) processing (Naglieri et al. 1983; Willis 1985). The successive (or sequential) processing requires the maintenance of the temporal order of input of information (Naglieri et al. 1983). An example of successive (or sequential) processing might include learning to read using a phonetic approach (Gunnison et al. 1982). That is not to say that reading cannot be accomplished by simultaneous processing such as the whole word approach, but rather with a phonetic approach, successive (or sequential) processing is more efficient.

Hemispheric Specialization and Simultaneous and Successive Cognitive Processes

Different cerebral hemispheres are thought to be more efficient with either simultaneous or successive processing (Naglieri et al. 1983). The left cerebral hemisphere may be more efficient in performing linguistic, serial, and analytic tasks. The right hemisphere may be more efficient in performing visual-spatial and gestalt-holistic tasks (Bever 1975; Bogen 1969; Dean and Reynolds 1997; Gazzaniga 1970; Harnad et al. 1977; Kinsbourne 1978, 1997; Naglieri et al. 1983; Schwartz et al. 1975; Segalowitz and Gruber 1977; Willis 1985). Modes of information processing appear likely related to hypothesized differences in cerebral hemispheric processing, and as earlier mentioned, the advantage is that one cerebral hemisphere may be more efficient in processing particular stimuli, but that does not mean that the other cerebral hemisphere cannot also process that same stimuli but rather a relative degree of efficiency may be lost. For example, there are persons who read using a whole word rather than a phonetic approach. Indeed, not all languages are phonetically based, so whole word approaches are essential in some languages. Utilization of specific cognitive processing modes may optimize the efficiency of these hemispheric brain functions.

Cerebral hemispheric asymmetries of functioning, as previously pointed out, are relative preferences for *process-specific* strategies rather than *stimulus-specific strategies*. The mode of higher-level cognitive processing for task performance depends on multiple factors which include, but are not limited to, specific task demands, level of attention required for the task, individual cognitive abilities, genetics, and cultural traditions (Cumming and Rodda 1985; Hall et al. 1988; McCallum and Merritt 1983; Watters and English 1995; Willis 1985). The need for specific types of manipulation of stimuli can also be a reason for selection of a specific hemispheric (e.g., Dean 1984; Grimshaw 1998; Mateer et al. 1984; Obrzut et al. 1985; Ornstein et al. 1980; Piccirilli et al. 1991; Tous et al. 1995).

Hemisphericity and Cognitive Processing

Reynolds (1981) conceptualized hemisphericity as preference for cognitive information-processing style independent of cerebral dominance. Hemisphericity can be defined as the tendency of an individual to rely differentially on the higher-level information-processing style of one cerebral hemisphere (Reynolds 1981). Previous research appears to be essentially consistent with hemispheric specialization (Dean and Reynolds 1997). Optimal higher-level cognitive functioning may require utilization of both modes of information and also being able to shift the cognitive information-processing mode in response to multiple factors (Gazzaniga 1974, 1975). At the same time, dysfunctional hemisphericity may impede optimal higher-level cognitive functioning (Newell and Rugel 1981; Roubinek et al. 1987). Research, over many years, has demonstrated that identifying the preferred mode of cognitive information processing (hemisphericity) may be advantageous in terms of addressing and remediating academic learning problems (Faust et al. 1993; Gunnison et al. 1982; Paquette et al. 1996; Roubinek et al. 1987; Sonnier 1992; Sonnier and Goldsmith 1985). Research on intelligence reviewed thus far has focused on intelligence as a single factor *g*, a brain-based behavior model, and different modes of cognitive information processing.

Halstead's Theory of Biological Intelligence

Simply put, theoretical interest in the human mental abilities subserved by the human frontal lobes was the focus of Ward Halstead's research program (Halstead 1947). This research interest also included the concept of intelligence (Halstead 1947). It might be noted that Boring (1930) has considered the concept of intelligence as what intelligence tests measured which is, of course, tautological.

Halstead accepted Boring's definition of intelligence as psychometric intelligence which was postulated to be what was measured by the intelligence tests (Halstead (1947). In contrast, however, Halstead also conceptualized a type of intelligence that was different from psychometric intelligence (Halstead 1947).

Biological intelligence as conceptualized by Halstead (1947) was human adaptive abilities as subserved by an intact uninjured brain, in other words, human adaptive abilities that were significantly impaired following brain damage.

The concept of biological intelligence was hypothesized in response to perceived limitations of intelligence tests. Halstead (1947) observed that in many cases, patients who had brain injuries were still able to score well on intelligence tests despite clear brain damage and significant adaptive behavior problems in daily living. Halstead (1947) conceptualized that there was an additional brain-based latent construct that was sensitive to human adaptive abilities but poorly evaluated by intelligence tests. In other words, Halstead agreed with David Wechsler's definition (1944) of intelligence as "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment" (p. 3) but found then contemporary intelligence tests inadequate to satisfy David Wechsler's definition (1944) of intelligence and sought to find a brain-based latent construct that would better satisfy Wechsler's definition (1944).

In order to research the latent construct of biological intelligence, Halstead (1947) established an experimental brain-behavior research program at the University of Chicago Medical School focused on studying the biological integrity of the human brain (Horton and Wedding 1984). Halstead (1947) developed a number of sensitive measures to the behavioral deficits of brain-injured persons which Halstead postulated represented the abstract concept of biological intelligence and were distinct from psychometric intelligence. Of particular interest is Halstead's factor analytic attempt to identify the aspects of higher cognitive functions that were involved in biological intelligence (Halstead 1947).

Halstead (1947) extracted four basic factors of biological intelligence, and these factors are described below:

C, *the integrative field factor*. The ability to adapt to new situations and to integrate new information was postulated as the integrative field factor (Reitan 1994). Tests that had loadings on factor C included the Halstead Category Test, the Henmon-Nelson Tests of Mental Ability, the Speech-Sounds Perception Test, the Halstead Finger Oscillation Test, and the Halstead Time-Sense Test (Halstead 1947).

A, *the abstraction factor*. The ability to draw meaning from a series of events or to hold in mind abstract nonverbal ideas without the use of past experience. Tests that had loadings on factor A included the Carlo Hollow-Square Performance Test for Intelligence, the Halstead Category Test, the Halstead Tactual Performance Test (memory component), and the Halstead Tactual Performance Test (localization component) (Halstead 1947).

P, *the power factor*. The reserve power available to an amplifier not already functioning at peak ability was postulated to be the power factor. Tests that had loadings on factor P included the Halstead Flicker-Fusion Test, the Halstead Tactual Performance Test (recall component), the Halstead Dynamic Visual Field Test (central form), and the Halstead Dynamic Visual Field Test (central color) (Halstead 1947).

D, *the directional factor*. An attentional component. Tests that had loadings on factor D included the Halstead Tactual Performance Test (speed component) and the Halstead Dynamic Visual Field Test (peripheral component) (Halstead 1947). The first three factors, C, A, and P, were interpreted as process factors of biological intelligence, and D was interpreted as the factor through which expressions of factors C, A, and P were directed.

It is noteworthy that both factors C and A had had significant loadings from intelligence tests but the intelligence tests loaded on different factors. The tests measuring the four factors significantly differentiated between individuals with documented head injury and individuals with no documented history of head injury (Halstead 1947).

Moreover, an average of the measures (the Halstead Impairment Index) was the best measure in differentiating these individuals (Halstead 1947). Unfortunately, as the concept of biological intelligence was postulated to be related to the integrity of the frontal lobes, subsequent experimental research studies couldn't cross-validate a relationship between the frontal lobes and HII (Reitan 1975), thus failing to confirm the concept of biological intelligence. In addition, as previously noted, tests of intelligence did load on factors extracted from Halstead's tests so the latent construct biological intelligence appeared to overlap rather than be orthogonal to psychometric intelligence. Interestingly, Halstead's tests were better able to differentiate brain damaged from normal subjects than intelligence tests alone but exactly why remains elusive.

Reitan (1994) had validated a modified and augmented neuropsychological test battery based on Halstead's tests as a core to improve diagnostic accuracy (Hevern 1980; Reed 1985; Swiercinsky 1979). It is noteworthy that formal intelligence testing has always included as an integral portion of Reitan's comprehensive clinical neuropsychological test battery, in addition to Halstead's core tests and a number of additional test procedures added to assess brain areas not related to intelligence tests. It is noteworthy that Halstead's factor structure (Horton and Wedding 1984) was very similar to the factor structure found for the age-appropriate Wechsler scales (Kamphaus 2001; Kaufman 1994). Basically, the contemporary Wechsler scales have a factor structure which consists of verbal comprehension, perceptual organization, working memory, and processing speed factors. In other words, very similar to the four factors identified by Halstead (1947).

Subsequent factor analysis studies with neuropsychological tests have produced comparable results. A few examples are cited. A study with adult neuropsychiatric patients (Fowler et al. 1988) extracted five factors (verbal comprehension, perceptual organization, sensory attention, primary motor, and tactual-spatial abilities). Also, a study of children aged 9–14 (Brooks et al. 1989) extracted 4 factors (simple motor, tactile kinesthesis, memory/attention, and nonverbal visual-spatial memory). In addition, a study with children aged 9–12 (Francis et al. 1992) extracted 5 factors (simple motor skill, complex visual-spatial relations, simple spatial motor operations, motor steadiness, and speeded motor sequencing). Moreover, a study with younger children aged 5–7 (Foxcroft 1989) extracted six factors (analytic-synthetic visual motor ability, perceptual organization, cross-modality motoric efficiency, directed motor speed, patterned critical discrimination, and strength). Briefly put, multiple factor analysis studies appear to extract factors which are relatively similar to factors found from the age-appropriate Wechsler scales (Kamphaus 2001; Kaufman 1994).

Therefore, conceptualizations of intelligence that are consistent with David Wechsler's (1944) definition of intelligence as "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment" might be seen as loading on the above factors. The relationship with *g* and the above factors needs to be reconciled. A possible answer, however, will be addressed in the next sections.

Contemporary Wechsler Scales of Intelligence

There are three contemporary Wechsler intelligence scales designed to assess adults, school-aged children, and children in preschool and primary grades. They include the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV) (Wechsler 2008), the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV) (Wechsler 2003), and the Wechsler Preschool and Primary Scale of Intelligence – Third Edition (WPPSI-III) (Wechsler 2002) and all allow examiners to compute full-scale IQs. The WAIS-IV is designed to assess from ages 16 to 90 and 11 months, the WISC-IV is designed to assess from ages 6 to 16 and 11 months, and the WPPSI-III is designed to assess from age 2 years and 6 months to age 7 years and 3 months. In terms of factor structures as recommended for clinical interpretation in the technical and interpretation manuals, there is a great deal of similarity. For the WAIS-IV (Wechsler 2008), the recommended factor structure forms the bases for the Verbal Comprehension Index, the Perceptual Reasoning Index, the Working Memory Index, and the Processing Speed Index. Similarly for the WISC-IV (Wechsler 2003), the recommended factor structure includes again the Verbal Comprehension Index, the Perceptual Reasoning Index, the Working Memory Index, and the Processing Speed Index. Essentially both the WAIS-IV and WISC-IV are reported to have the same factor structure. For the WPPSI-III (Wechsler 2002), however, there are differences depending on the age of the child. Essentially from ages 2 years and 6 months to 3 years and 11 months, a two-factor model is recommended with a verbal intelligence quotient (VIQ) and performance intelligence quotient (PIQ). From age 4 to 7 years and 3 months, a three-factor model is recommended with a verbal intelligence quotient (VIQ) and performance intelligence quotient (PIC) and a processing speed quotient (PSQ). In summary, all of the Wechsler scales of intelligence show verbal and performance factors at every age level. For the WPPSI-III, a two-factor model of verbal and performance is preferred from ages 2 years and 6 months to 3 years and 11 months; in a three-factor model of verbal performance and processing, speed is preferred from ages 4 to 7 years and 3 months. For the WISC-IV and WAIS-IV, four-factor models (verbal comprehension, perceptual reasoning, working memory, and processing speed) are preferred. As earlier mentioned, all of the Wechsler scales allow the computation of a full-scale IQ. Recent research studies (Canivez and Watkins 2010) supported the WAIS-IV as a measure of general intelligence but noted the remaining factor structure accounted for small portions of total and common variance. Benson et al. (2010) suggested that a Cattell-Horn-Carroll (CHC) structure provides a better description of test performance with abilities that include crystallized ability (Gc), fluid reasoning (Gf), visual processing (Gv), short-term memory (Gsm), and processing speed (Gs). Moreover, Weiss et al. (2013a) found that either a four- or five-factor structure fits the data, but a five-factor structure was a better fit with a quantitative reasoning (RQ) factor included.

For the WISC-IV, Keith et al. (2006) suggested the scoring structure was not supported and the Cattell-Horn-Carroll (CHC) theory was a better fit, but Watkins (2010) found the four first-order factors as suggested by the WISC-IV test manual (Wechsler 2003). More recently, Weiss et al. (2013) found that either a four- or five-factor structure fits the data and both were suitable, and the five-factor model included inductive reasoning (IR).

A common concern was that the *g* factor was the majority of the variance and other factors were quite small.

Carroll's Theory of Intelligence

Carroll's (1993) three-stratum theory of intelligence has averred that the latent traits tapped by intelligence tests are independent of the specific test battery. Carroll (1993) has postulated that numerous mental ability tests measured the same abilities which Carroll labeled crystallized, visual-perceptual, and memory abilities. In interpreting extant research findings, Carroll (1993) has proposed there are three strata of intelligence. An important feature is the reconciliation of previous research results related to the assessment of human intelligence by combining the Cattell-Horn notion of crystallized G (Gc) and fluid G (gf) with the Carroll paradigm into the Cattell-Horn-Carroll (CHC) theory (McGrew 2009). In Carroll's (1993) theory, the third stratum is unitary or, put another way, is composed of one construct only, *g* as previously described. Multiple studies of human intelligence have found that *g* accounts for the major portion of variance assessed by intelligence test batteries. Similarly, intelligence tests are strong and consistent predictors of very important social outcomes, such as academic achievement (Binet and children in the Paris Public Schools) and occupational performance (Wechsler and officer candidates in the US Army in World War I and in addition the SATs, GREs, etc.). The predictive ability is directly related to the amount of *g* measured by the intelligence test. Simply put, intelligence tests with greater amounts of *g* are significantly better predictors of important outcomes in society than are intelligence tests with lower amounts of *g*. Clearly intelligence tests with large amounts of *g* have important purposes in society, especially in terms of prediction of success in academic and occupational settings. It is a conundrum that while the psychometric concept of *g* has proven useful in society for over a century, the full understanding of the latent concept of *g* remains elusive and is not yet completely understood by psychologists even after a century of research and clinical application. The CHC theory posits various types of *g* such as Gc and Gf among others.

Carroll's (1993) second stratum of traits is composed of combinations of stratum one measures and second-stratum measures that combine to form the third stratum. Typically, stratum one measures are more specific traits of interest. Stratum one measures are combined to become stratum two measures and result in enhanced measurement of complex higher-level cognitive traits such as verbal and nonverbal intelligence. Similarly, stratum two measures are then in a hieratical fashion combined to allow for the measurement of a complex stratum three trait, such as the latent construct of intelligence or *g*. Concepts such as fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad auditory perception, and processing speed are examples of second-stratum traits (Carroll 1993). Multiple research studies appear to suggest second-stratum traits can be ranked in terms of their abilities to assess *g* (Kamphaus 2001). Second-stratum traits which involve reasoning abilities are better measures of *g*. Examples of second-stratum traits that involve abstraction abilities might be seen as general sequential reasoning, induction, deduction, syllogisms, series tasks, matrix reasoning, analogies and quantitative reasoning, etc. (Carroll 1993). An example of a contemporary intelligence test that uses the CHC theory as a basis is the Reynolds Intellectual Assessment Scale (RIAS) which will be described in the next section.

Reynolds Intellectual Assessment Scale (RIAS)

The Reynolds Intellectual Assessment Scale (RIAS) (Reynolds and Kamphaus 2003) follows the more contemporary Carroll (1993) theoretical model of intelligence model. The RIAS has demonstrated impressive evidence for its interpretation as a measure of intelligence (i.e., validity) (Beaujean et al. 2010) as well as being time efficient, user-friendly for administration and scoring, and not having a disparate impact when used to assess members of minority groups, different genders, or groups of clinical patients. The RIAS proposed two-factor structure (verbal intelligence and nonverbal intelligence) has been cross-validated a number of times (Nelson et al. 2007; Dombrowski et al. 2009; Nelson and Ganivez 2012).

Discussion: Common and Variable Aspects of Intelligence

The concept of intelligence appears clearly related to the biological integrity of the brain (Luria 1973). Perhaps not solely to the frontal lobes alone (Reynolds and Horton 2006), but clearly intelligence is related to optimal human brain functioning (Reitan 1994). This chapter has demonstrated that the concept of intelligence can be conceptualized in multiple ways. Carroll's (1993) three-stratum theory of intelligence and the CHC model has found that latent mental traits are test battery independent and numerous tests measured the same latent mental traits tapped by intelligence tests. Multiple research studies have found *g* accounts for the major portion of variance assessed by intelligence test batteries. Also, Carroll's (1993) second stratum consists of higher-level traits such as verbal and nonverbal intelligence (Reynolds and Kamphaus 2003) that are assessed by combinations of stratum one measures. Stratum one measures are typically single subtests that measure a trait of interest and can be combined to form stratum two measures and measure higher-level cognitive abilities such as fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad auditory perception, and processing speed. In turn, stratum two measures are combined into a complex stratum three trait, such as general intelligence as conceptualized as *g*.

Therefore, intelligence can be conceptualized on multiple theoretical levels. Intelligence can be seen as represented by a single score that has impressive predictive abilities, different cognitive processing modes that have implications for higher cognitive functioning and multiple more specific higher-level cognitive ability factors that represent less comprehensive important cognitive skills (Carroll 1993). Relative to the most appropriate conceptualization of the latent trait of intelligence, the two-factor model exemplified by the RIAS appears the right choice. As previously mentioned, it appears the greater amount of *g* accounted for the theoretically more appropriate measure of intelligence. The contemporary intelligence test that maximizes the utilization of *g* is the RIAS. It should be recalled that for the WPPSI-III, the younger age has only two factors and for the older age of the WPPSI-III and for the WISC-IV and WAIS-IV, the later-appearing factors such as working memory and processing speed generally account for lesser amounts of *g*. In the earlier discussion of Halstead's factor analysis of Halstead's neuropsychological tests, it was noted that the first two factors included intelligence tests of the day. Indeed, the WISC-IV and WAIS-IV (Wechsler 2008) now have a measure known as the General Ability Index (GAI) which is a composite score of the three subtests that make up the VCI and the three subtests that make up the PRI. The GAI is proposed to be used when working memory and processing speed measures may have been impaired due to neuropsychological problems (Wechsler 2008).

In other words, the two-factor solution is a superior measure than FSIQ with these clinical groups (Wechsler 2008). Thinking back to the earlier mentioned studies of factor analyses of neuropsychological test batteries, it would seem that human mental abilities in excess of a two-factor solution such as the RIAS and GAI might be better characterized as neuropsychological abilities rather than intelligence (Reynolds and Kamphaus 2003). Simply put, the common structure of intelligence is composed of *g*, and the most *g* loaded two factors (verbal intelligence and nonverbal intelligence) are the best approximation of the latent construct of intelligence and also are the most appropriate basis for a contemporary comprehensive intelligence test.

Development of the various and common aspects of intelligence involve brain mechanisms assisted by cultural-historical experiences, as suggested by Luria (1966, 1973). Intelligence has been conceptualized as certainly influenced by the person's environmental history (i.e., for a discussion of Luria's Cultural-Historical Theory, see Horton 1987, Reynolds 1981) but also with genetic influences mediating the functional development of the various anatomical structures of the brain. Intelligence appears related to an individual's ability to adapt to various life circumstances (Pallier et al. 2000). Further developments of the theoretical foundations of intelligence appear likely to continue to elucidate how the human brain carries out higher-order cognitive functioning. A number of excellent new measures of intelligence have been developed in the past two decades (e.g., Kaufman and Kaufman 1983; Naglieri et al. 2013; Reynolds and Kamphaus 2003), but additional research related to the elaboration of the latent concept of intelligence is needed. The pace of new knowledge is expected to increase and more differentiated and complex understanding of the latent concept of human intelligence is expected.

References

Aluja-Fabregat, A., Colom, R., Abad, F., & Juan-Espinosa, M. (2000). Sex differences in general intelligence defined as *g* among young adolescents. *Personality and Individual Differences, 28*, 813–820.

Ashman, A. F., & Das, J. P. (1980). Relation between planning and simultaneous-successive processing. *Perceptual and Motor Skills, 51*, 371–382.

Beaujean, A. A., Firmim, M. W., Michonski, J. D., & Berry, T. (2010). A multitrait-multimethod examination of the Reynolds Intellectual Assessment Scales in a college sample. *Assessment, 17*(3), 347–360.

Benson, N., Hulac, D. M., & Kranzler, J. H. (2010). Independent examination of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV): what does the WAIS-IV measure? *Psychological Assessment, 22*(1), 121–130.

Bever, T. G. (1975). Cerebral asymmetries in humans are due to the differentiation of two incompatible processes: Holistic and analytic. In D. Aaronson & R. Reiber (Eds.), *Developmental psycholinguistics and communication disorders*. New York: New York Academy of Sciences.

Binet, A., & Simon, T. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'Année Psychologique, 11*, 191–244.

Binet, A., & Simon, T. (1908). The development of intelligence in the child. *L'Année Psychologique, 14*, 1–90.

Bogen, J. E. (1969). The other side of the brain: Parts I, II, and III. *Bulletin of the Los Angeles Neurological Society*, *34*, 73–105, 135–162, 191–203.

Boring, E. G. (1929). *History of experimental psychology*. New York: Genesis Publishing.

Brand, C. (1996). Doing something about *g*. *Intelligence, 22*, 311–326.

Brooks, D. A., Dean, R. S., & Gray, J. W. (1989). HRNB factors with simultaneous-successive marker variables. *International Journal of Neuroscience, 46*(3–4), 157–160.

Canivez, G. L., & Watkins, M. W. (2010). Investigation of the factor structure of the Wechsler Adult intelligence Scale-Fourth Edition (WAIS-IV): Exploratory and higher order factor analysis. *Psychological Assessment, 22*(4), 827–836.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Cumming, C. E., & Rodda, M. (1985). The effects of auditory deprivation on successive processing. *Canadian Journal of Behavioural Science, 17*, 232–245.

Das, J. P., Kirby, J. R., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processes*. New York: Academic.

Dean, R. S. (1984). Functional lateralization of the brain. *Journal of Special Education, 18*(3), 239–256.

Dean, R. S., & Reynolds, C. R. (1997). Cognitive processing and self-report of lateral preference. *Neuropsychology Review, 7*, 127–142.

Detterman, D. K. (1982). Does "g" exist? *Intelligence, 6*, 99–108.

Dombrowski, S. C., Watkins, M. W., & Brogan, M. J. (2009). An exploratory investigation of the factor structure of the Reynolds Intellectual Assessment Scales (RIAS). *Journal of Psychoeducational Assessment, 27*(6), 494–507.

Faust, M., Kravetz, S., & Babkoff, H. (1993). Hemisphericity and top-down processing of language. *Brain and Language, 44*, 1–18.

Fowler, P. C., Zillmer, E., & Newman, A. C. (1988). A multifactor model of the Halstead-Reitan neuropsychological test battery and its relationship to cognitive status and psychiatric diagnosis. *Journal of Clinical Psychology, 44*(6), 898–906.

Foxcroft, C. D. (1989). Factor analysis of the Reitan-Indiana neuropsychological test battery. *Perceptual and Motor Skills, 69*(3 pt. 2), 1303–1313.

Francis, D. J., Fletcher, J. M., Rourke, B. P., & York, M. J. (1992). A five-factor model for the motor, psychomotor and visual-spatial tests used in the neuropsychological assessment of children. *Journal of Clinical and Experimental Neuropsychology, 14*(4), 625–637.

Gazzaniga, M. S. (1970). *The bisected brain*. New York: Appleton.

Gazzaniga, M. S. (1974). Cerebral dominance viewed as a decision system. In S. Dimond & J. Beaumont (Eds.), *Hemisphere functions in the human brain*. London: Halstead Press.

Gazzaniga, M. S. (1975). Recent research on hemispheric lateralization of the human brain: Review of the split-brain. *UCLA Educator*, *May*, 9–12.

Golden, C. J. (1978). *Diagnosis and rehabilitation in clinical neuropsychology*. Springfield: Charles C. Thomas.

Golden, C. J. (1987). *Luria-Nebraska neuropsychological battery children's revision*. Los Angeles: Western Psychological Services.

Golden, C. J., Purisch, A. D., & Hammeke, T. A. (1979). *The Luria-Nebraska neuropsychological test battery: A manual for clinical and experimental uses*. Lincoln: The University of Nebraska Press.

Grimshaw, G. M. (1998). Integration and interference in the cerebral hemispheres: Relations with hemispheric specialization. *Brain and Cognition, 36*, 108–127.

Gunnison, J., Kaufman, N. L., & Kaufman, A. S. (1982). Reading remediation based on sequential and simultaneous processing. *Academic Therapy, 17*, 297–306.

Hall, C. W., Gregory, G., Billinger, E., & Fisher, T. (1988). Field independence and simultaneous processing in preschool children. *Perceptual and Motor Skills, 66*, 891–897.

Halstead, W. C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.

Harmony, T. (1997). Psychophysiological evaluation of neuropsychological disorders in children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 356–370). New York: Plenum Press.

Harnad, S., Doty, R. W., Goldstein, L., Jaynes, J., & Krauthamer, G. (Eds.). (1977). *Lateralization in the nervous system*. New York: Academic.

Hevern, V. W. (1980). Recent validity studies of the Halstead-Reitan approach to clinical neuropsychological assessment: A critical review. *Clinical Neuropsychology, 2*, 49–61.

Horton, A. M., Jr. (1987). Luria's contributions to clinical and behavioral neuropsychology. *Neuropsychology, 1*(2), 39–44.

Horton, A. M., Jr., & Wedding, D. (1984). *Clinical and behavioral neuropsychology*. New York: Praeger.

Jensen, A. R. (1978, September). *"g":Outmoded concept or unconquered frontier?* Invited address at the annual meeting of the American Psychological Association, New York.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.

Kamphaus, R. W. (2001). *Clinical assessment of child and adolescent intelligence* (2nd ed.). Boston: Allyn and Bacon.

Kane, H. D. (2000). A secular decline in Spearman's *g*: Evidence from the WAIS, WAIS-R and WAIS-III. *Personality and Individual Differences, 29*, 561–566.

Kaufman, A. S. (1979b). Cerebral specialization and intelligence testing. *Journal of Research and Development in Education, 12*, 96–107.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.

Kaufman, A. S., & Kaufman, N. L. (1983). *Administration and scoring manual for the Kaufman assessment battery for children*. Circle Pines: American Guidance Service.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order multi-sampled confirmatory factor analysis of the Wechsler intelligence scale children-fourth edition: What does it measure? *School Psychology Review, 35*(1), 108–127.

Kinsbourne, M. (1978). Biological determinants of functional bisymmetry an asymmetry. In M. Kinsbourne (Ed.), *Asymmetrical function of the brain*. Cambridge: Cambridge University Press.

Kinsbourne, M. (1997). Mechanisms and development of cerebral lateralization in children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 102–119). New York: Plenum Press.

Kirby, J. R., & Das, J. P. (1977). Reading achievement, IQ, and simultaneous-successive processing. *Journal of Educational Psychology, 69*, 564–570.

Languis, M. L., & Miller, D. C. (1992). Luria's theory of brain functioning: A model for research in cognitive psychophysiology. *Educational Psychologist, 27*, 493–511.

Luborsky, L., Auerbach, A. H., Chandler, M., Cohen, J., & Bachrach, H. M. (1971). Factors influencing the outcome of psychotherapy: A review of quantitative research. *Psychological Bulletin, 75*, 145–185.

Luria, A. R. (1964). Neuropsychology in the local diagnosis of brain damage. *Cortex, 1*, 3–18.

Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.

Luria, A. R. (1973). *The working brain*. London: Penguin.

Mateer, C. A., Rapport, R. L., & Kettrick, C. (1984). Cerebral organization of oral and signed language responses: Case study evidence from amytal and cortical stimulation studies. *Brain and Language, 21*, 123–135.

McCallum, R. S., & Merritt, F. M. (1983). Simultaneous-successive processing among college students. *Journal of Psychoeducational Assessment, 1*, 85–93.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1–10.

Naglieri, J. A., Kamphaus, R. W., & Kaufman, A. S. (1983). The Luria-Das simultaneous-successive model applied to the WISC-R. *Journal of Psychoeducational Assessment, 1*, 25–34.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2013). *Cognitive assessment system*(2nd ed.). Austin, TX: ProEd.

Nelson, J. M., & Canivez, G. L. (2012). Examination of the structural, convergent, and incremental validity of the Reynolds Intellectual Assessment scale (RIAS) with a clinical simple. *Psychological Assessment, 24*(1), 129–140.

Nelson, J. M., Canivez, G. L., Lindstrom, W., & Hatt, C. V. (2007). Higher-order exploratory factor analysis of the Reynolds intellectual assessment scales with a referred sample. *Journal of School Psychology, 45*(4), 439–456.

Newell, D., & Rugel, R. P. (1981). Hemispheric specialization in normal and disabled readers. *Journal of Learning Disabilities, 14*, 296–298.

Obrzut, J. E., & Obrzut, A. (1982). Neuropsychological perspectives in pupil services: Practical application of Luria's model. *Journal of Research and Development in Education, 15*, 38–47.

Obrzut, J. E., Obrzut, A., Bryden, M. P., & Bartels, S. G. (1985). Information processing and speech lateralization in learning-disabled children. *Brain and Language, 25*, 87–101.

Ornstein, R., Johnstone, J., Herron, J., & Swencionis, C. (1980). Differential right hemisphere engagement in visuospatial tasks. *Neuropsychologia, 18*, 49–64.

Pallier, G., Roberts, R. D., & Stankov, L. (2000). Biological versus psychometric intelligence: Halstead's (1947) distinction revisted. *Archives of Clinical Neuropsychology, 15*, 205–226.

Paquette, C., Tosoni, C., Lassonde, M., & Peretz, I. (1996). Atypical hemispheric specialization in intellectual deficiency. *Brain and Language, 52*, 474–483.

Piccirilli, M., D'Alessandro, P., Mazzi, P., Sciarma, T., & Testa, A. (1991). Cerebral organization for language in Down's syndrome patients. *Cortex, 27*, 41–47.

Reed, J. (1985). The contributions of Ward Halstead, Ralph Reitan and their associates. *International Journal of Neuroscience, 25*, 289–293.

Reitan, R. M. (1975). Assessment of brain-behavior relationships. In P. McReynolds (Ed.), *Advances in psychological assessment: Vol. III*. San Francisco: Jossey-Bass.

Reitan, R. M. (1994). Ward Halstead's contributions to neuropsychology and the Halstead-Reitan neuropsychological test battery. *Journal of Clinical Psychology, 50*, 47–69.

Reynolds, C. R. (1981). The neuropsychological basis of intelligence. In G. Hynd & J. Obrzut (Eds.), *Neuropsychological assessment and the school-aged child*. New York: Grune & Stratton.

Reynolds, C. R., & French, C. (2003). Neuropsychological basis of intelligence: Revisited. In A. M. Horton Jr. & L. C. Hartlage (Eds.), *Handbook of forensic neuropsychology*. New York: Springer.

Reynolds, C. R., & Horton, A. M., Jr. (2006). *Test of verbal conceptualization and fluency: Examiner's manual*. Austin: PRO-ED.

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds intellectual assessment scales*. Odessa: PAR.

Roubinek, D. L., Bell, M. L., & Cates, L. A. (1987). Brain hemispheric preference of intellectually gifted children. *Roeper Review, 10*, 120–122.

Schwartz, G. E., Davidson, R. J., & Maer, F. (1975). Right hemisphere lateralization for emotion in the human brain: Interactions with cognition. *Science, 190*, 286–288.

Segalowitz, S. J., & Gruber, F. A. (Eds.). (1977). *Language development and neurological theory*. New York: Academic.

Shure, G. H., & Halstead, W. C. (1958). Cerebral localization of Intellectual processes. *Psychological Monographs: General and Applied, 72*(12), 1–40.

Sonnier, I. L. (1992). Hemisphericity as a key to understanding individual differences. In I. L. Sonnier (Ed.), *Hemisphericity as a key to understanding individual differences* (pp. 6–8). Springfield: Charles C. Thomas.

Sonnier, I. L., & Goldsmith, J. (1985). The nature of human brain hemispheres: The basis for some individual differences. In I. L. Sonnier (Ed.), *Methods and techniques of holistic education* (pp. 17–25). Springfield: Charles C. Thomas.

Swiercinsky, D. P. (1979). Factorial pattern description and comparison of functional abilities in neuropsychological assessment. *Perceptual and Motor Skills, 48*, 231–241.

Tous, J. M., Fusté, A., & Vidal, J. (1995). Hemispheric specialization and individual differences in cognitive processing. *Personality and Individual Differences, 19*, 463–470.

Travers, R. M. W. (1977). *Essentials of learning* (4th ed.). New York: Macmillan.

Vernon, P. A. (1998). From the cognitive to the biological: A sketch of Arthur Jensen's contributions to the study of *g*. *Intelligence, 26*, 267–271.

Watkins, M. W. (2010). Structure of the Wechsler intelligence scale for children-fourth edition among national sample of referred students. *Psychological Assessment, 22*(4), 782–787.

Watters, J. J., & English, L. D. (1995). Children's application of simultaneous and successive processing in inductive and deductive reasoning problems:

Implications for developing scientific reasoning skills. *Journal of Research in Science Teaching, 32*, 699–714.
Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.
Wechsler, D. (1944). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.
Wechsler, D. (1952). *The range of human capacities*. Baltimore: Williams & Wilkins.
Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence-third edition: Technical interpretive manual*. San Antonio: The Psychological Corporation.
Wechsler, D. (2003). *Wechsler intelligence scale for children-fourth edition: Technical and interpretive manual*. San Antonio: The Psychological Corporation.
Wechsler, D. (2008). *Wechsler adult intelligence scale-fourth edition: Technical interpretive manual*. San Antonio: NCS Pearson.
Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013a). WAIS-IV and clinical validation of the four- and five factor interpretive approaches. *Journal of Psychoeducational Assessment, 31*(2), 94–113.
Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013b). WISC-IV and clinical validation of the four- and five factor interpretive approaches. *Journal of Psychoeducational Assessment, 31*(2), 114–131.
Willis, W. G. (1985). Successive and simultaneous processing: A note on interpretation. *Journal of Psychoeducational Assessment, 4*, 343–346.

Current Concepts in the Assessment of Emotional Intelligence

24

Steven J. Stein and Justin M. Deonarine

All learning has an emotional base.

Plato

Emotional intelligence (EI) refers to the ability to perceive, manage, and evaluate emotions. Some researchers suggest that emotional intelligence can be learned and strengthened, while others claim it is an inborn characteristic.

The concept of emotional intelligence has brought new depth to the understanding of human intelligence; it enhances the ability to evaluate one's general or overall intelligence. Unlike cognitive intelligence, emotional intelligence has been difficult to define. Broadly speaking, emotional intelligence addresses the emotional, personal, social, and survival dimensions of intelligence, which are often more important for daily functioning than the more traditional cognitive aspects of intelligence. Emotional intelligence is concerned with understanding oneself and others, relating to people, and adapting to and coping with the immediate surroundings in order to be more successful in dealing with environmental demands. Emotional intelligence is tactical (immediate functioning), while cognitive intelligence is strategic (long-term capacity). Emotional intelligence helps to predict success, because it reflects how a person applies knowledge to the immediate situation. In a way, to measure emotional intelligence is to measure one's "common sense" and ability to get along in the world.

Ever since the first peer-reviewed publication on emotional intelligence appeared (Salovey and Mayer 1990) and its later popularization in the trade publications (e.g., Goleman 1995), there has been controversy over the concept. A number of academics have attacked the models presented for their definitional fuzziness, lack of empirical data, inconsistencies of the various measures developed, and so on. Many of the same criticisms applied to emotional intelligence could easily be applied to cognitive intelligence, such as the fact that there is still no uniformly accepted single theory after over 100 years of research. Regardless of their claims, the popularity, use, and acceptance of the term "emotional intelligence" have now been firmly established. A Google search on "emotional intelligence" currently brings up over 25,000,000 entries.

In addition, the emerging field of emotional intelligence has been attacked for its disproportionate emphasis on application before validation. This has created several gaps in the field that have been pointed out in a number of critical commentaries (e.g., Landy 2005; Zeidner et al. 2004; Murphy 2006). Much of the skepticism surrounding the applications of the EI construct comes from issues that are directly related to its assessment strategies. One issue concerns the methodological treatment of the construct's conceptual multidimensionality.

S.J. Stein, Ph.D. (✉) • J.M. Deonarine, M.A.
Multi-Health Systems, 3770 Victoria Park Ave., Toronto ON M2H 3M6, ON, Canada
e-mail: ceo@mhs.com; justin.deonarine@mhs.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_24,

With roots going back to different theoretical traditions, the nature of EI has been elaborated in a number of comprehensive models comprising different dimensions and spanning multiple psychological systems, domains, and processes (Bar-On and Parker 2000). This theoretical complexity, however, cannot be found in the applied arena, where the practice of combining various EI dimensions into a single composite score and treating different models and measures of EI as mutually interchangeable or otherwise incompatible has been quite common (Zeidner et al. 2008).

Failure to account for the non-homogeneity of complex constructs is not only likely to complicate their true nature and functions but also can lead to inaccurate decisions and treatment strategies for individual test takers (McGrath 2005; Smith et al. 2009). Adopting a multidimensional approach toward the study and measurement of EI is clearly needed to remove the ambiguity of past research and move the EI field forward. Another major obstacle to progress in the field has been the proliferation and use of EI measures with inadequate or under researched psychometric properties (Zeidner et al. 2008).

Brief History of Emotional Intelligence

Many people, upon first hearing the term emotional intelligence, presume that it is an oxymoron. After all, how could one be emotional while at the same time be intelligent? Most likely, emotional intelligence is as old as time. In the 1870s, Charles Darwin published the first modern book on the role of emotional expression in survival and adaptation. The heavily illustrated book presents the thesis that emotions are universal—not only among cultures but across species. Additionally, emotions serve a function that aided in the survival of species over time. They served as an early warning system, alerting the individual (and possibly others in flocks) that danger was near and that a fight or flight response was required. Another example is the warning of a threat to an offspring where protection was required. Basically, emotions are a call to action. The ability to successfully rise to the call is one's way of successfully coping with one's environment—a component of most definitions of intelligence.

In addition, one's emotional expression might serve as a warning to others, such as to stay away when angry, or as a signal to the herd if an impending danger is sensed. An expression of sadness or hurt might signal others to help. One's ability to use these emotional signals, either for self-motivation, efficiently alerting others, or reading another's emotions, is identified as an intelligence as it varies across individuals and, once again, influences one's ability to adapt to the environment.

However, to gain a practical perspective, we will focus on the development of the concept of emotional intelligence or one's emotional quotient (EQ) as developed in the twentieth and twenty-first centuries. Back in the 1920s, the American psychologist Edward Thorndike wrote about a concept he called "social intelligence." He attempted to measure one's social skills and abilities through a paper and pencil test.

Later, the importance of "emotional factors" was recognized by David Wechsler, one of the fathers of IQ testing. In a rarely cited paper written in the 1940s, Wechsler urged that the "non-intellective aspects of general intelligence" be included in any "complete" measurement of intelligence. This paper also discussed what he called "affective" and "conative" abilities—basically, emotional and social intelligence—which he thought would prove critical to an overall view. Unfortunately, these factors were not included in Wechsler's IQ tests, and little attention was paid to them at the time (Wechsler 1940, 1943).

In 1948, another American researcher, R. W. Leeper, promoted the idea of "emotional thought," which he believed contributed to "logical thought." But few psychologists or educators pursued this line of questioning until more than 30 years later. (One notable exception was Albert Ellis, who, in 1955, began to explore what would become known as rational emotive behavior therapy—a process that involved teaching people to examine their emotions in a logical, thoughtful way.) In his 1976 book "The Shattered Mind" and his 1983 book “Frames of Mind: The Theory of Multiple Intelligences,” Howard Gardner wrote about the possibility of “multiple intelligences,” including what he called “intra-psychic capacities”—in essence, an aptitude for introspection—and “personal intelligence” Gardner (1976, 1983).

By this time, Reuven Bar-On was completing his dissertation in which he had contributed the phrase “emotional quotient” (EQ) as a measure of emotional intelligence based on his developing model and measure of this construct. The evolution of the Emotional Quotient Inventory (EQ-i) as an assessment of emotional intelligence began in 1980 with the independent development of a theoretically eclectic and multifactorial approach to operationally defining and measuring emotional intelligence. The seminal work of Bar-On (1988) was inspired by his work as a clinical psychologist, with the goal of answering the question, “Why do some people have better psychological well-being than others?” This question ultimately expanded into, “Why are some individuals better able to succeed in life than others?” These questions commanded a thorough review of the factors (emotional skills) thought to determine general success, in addition to maintaining positive emotional health. It soon became clear that the key to determining and predicting success is not cognitive intelligence alone, as many cognitively intelligent people flounder in life, while many less cognitively intelligent individuals succeed and prosper. During the 3-year period between 1983 and 1986 while he completed his doctoral studies in South Africa, Bar-On had directed his efforts at identifying the most important factors involved in coping with environmental demands, at which point his research revealed a nonsignificant relationship between cognitive intelligence and emotional intelligence.

Psychologists continued to challenge the conventional view of intelligence. John Mayer at the University of New Hampshire and Peter Salovey at Yale University concentrated their research efforts on the “emotional” aspect of intelligence (Mayer 1986; Mayer et al. 1988, 1990, 1991; Mayer and Salovey 1988, 1993, 1995; Mayer and Volanth 1985; Salovey and Birnbaum 1989; Salovey et al. 1991, 1993; Salovey and Mayer 1990; Salovey and Rodin 1985). In 1990, Mayer and Salovey would coin and formally define the term “emotional intelligence” as “the subset of social intelligence that involves the ability to monitor one’s own and others’ feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions.” Together with David Caruso they expanded upon Gardner’s approach and looked primarily at six components of emotional intelligence that are very similar to what Bar-On (1997) refers to as emotional self-awareness, assertiveness, empathy, interpersonal relationship, stress tolerance, and impulse control.

The release of Goleman’s (1995) “Emotional Intelligence: Why It Can Matter More Than IQ” served to popularize the construct of emotional intelligence, rendering it a hot topic among academics, coaches, consultants, psychologists, and the layperson alike, thereby paving the way for—and in some cases creating a demand for—a valid and reliable measure of EI. However, Goleman’s writing and “theories” of emotional intelligence have been heavily criticized by Mayer for going far beyond their original work, which he drew from (Mayer et al. 2004).

Finally, the extensive work of Bar-On culminated in 1997 with the release of the first psychometrically valid and reliable measure of emotional intelligence, the EQ-i (Bar-On 1997). Interest in the EQ-i steadily increased since its release in 1997. The energy created and sustained by it is evidenced by the number of publications applying or investigating the EQ-i, including books (13), trade publications (countless), dissertations (77), and peer-reviewed articles (71). The EQ-i also garnered attention in several peer-reviewed journal special issues including *Emotion*, *Psychological Inquiry*, and the *Journal of Organizational Behavior*. The EQ-i has captured the attention of researchers and practitioners worldwide. It can be employed in many ways and in a variety of settings.

The term emotional intelligence (EI) has significantly evolved since the first release of the EQ-i, enduring rigorous debate over how to define it, how to measure it, whether it can be developed, and whether or not it adds incremental value over and above personality and IQ. This intense scrutiny of EI helped to refine the construct; not only has EI endured, its utility is more deeply understood, and its use more widespread. As a result, the operational definition of emotional intelligence as it relates to the EQ-i 2.0 is "…a set of emotional and social skills that influence the way we perceive and express ourselves, develop and maintain social relationships, cope with challenges, and use emotional information in an effective and meaningful way."

Here is a brief history of events leading up to the current concept of emotional intelligence:

- 1930s – Edward Thorndike describes the concept of "social intelligence" as the ability to get along with other people.
- 1940s – David Wechsler suggests that affective components of intelligence may be essential to success in life.
- 1950s – Humanistic psychologists such as Abraham Maslow describe how people can build emotional strength.
- 1976 – Howard Gardner publishes *The Shattered Mind*, which introduces the concept of multiple intelligences. This book was followed by *Frames of Mind: The Theory of Multiple Intelligences* in 1983.
- 1985 – Reuven Bar-On introduced the term EQ (emotional quotient) in his doctoral dissertation in which he created a test to measure one's emotional quotient.
- 1990 – Psychologists Peter Salovey and John Mayer publish their landmark article, "Emotional Intelligence," in the journal *Imagination, Cognition, and Personality*.
- 1995 – The concept of emotional intelligence is popularized after publication of psychologist and *New York Times* science writer Daniel Goleman's book *Emotional Intelligence: Why It Can Matter More Than IQ*.

Models of Emotional Intelligence

Emotional Quotient Inventory (2.0)

Basically, Bar-On arrived at a way to capture emotional and social intelligence by dividing it into five general areas or realms and 15 subsections or scales. Based on updated research and the latest theories on emotional intelligence, the MHS team has created an updated model of emotional intelligence that is measured by the Emotional Quotient Inventory 2.0 (EQ-i 2.0). It now includes 16 subsections or scales as illustrated in Fig. 24.1.

The five realms and their scales include the following:

The self-perception realm concerns one's ability to know and manage oneself. It embraces emotional self-awareness, the ability to recognize feelings and why one feels that way and the impact one's emotions have on thoughts and actions of oneself and others; self-regard, the ability to recognize strengths and weaknesses and to feel good about oneself despite any weaknesses; and self-actualization, the ability persistently try to improve and pursue meaningful goals that lead to a richer life.

The self-expression realm deals with the way one faces the world. Emotional expression is the ability to express feelings both in words and non-verbally. Independence is the ability to be self-directed and self-controlled, to stand on one's own two feet; assertiveness is the ability to clearly express thoughts and beliefs, stand one's ground, and defend a position in a constructive way.

The interpersonal realm concerns "people skills"—the ability to interact and get along with others. It is composed of three scales. Empathy is the ability to recognize, understand, and appreciate what others may be feeling and thinking. It is the ability to view the world through another person's eyes. Social responsibility is the ability to be a cooperative and contributing member of one's social group and to society at large. Interpersonal relationships refer to the ability to forge and maintain relationships that are mutually beneficial and marked by give and take and a sense of trust and compassion.

The decision-making realm involves the ability to use emotions in the best way to help solve problems and make optimal choices. Its three scales are impulse control, the ability to resist or delay a temptation to act rashly; reality testing, the ability to see things as they actually are, rather than the way one wishes or fears they might be; and problem solving, the ability to find solutions

Fig. 24.1 Model of emotional intelligence (Reprinted with permission of Multi-Health Systems, Inc., Toronto, Canada. www.mhs.com)

to problems where emotions are involved using the right emotion at an optimum value.

The stress management realm concerns the ability to be flexible, tolerate stress, and control impulses. Its three scales are flexibility, the ability to adjust feelings, thoughts, and actions to changing, challenging, or unfamiliar conditions; stress tolerance, the ability to remain calm and focused, to constructively withstand adverse events and conflicting emotions without caving in; and optimism, the ability to maintain a realistically positive attitude, particularly in the face of adversity.

There is also an independent indicator of happiness. Happiness is the ability to feel satisfied with life, to enjoy oneself and others, and to experience zest and enthusiasm in a range of activities (Table 24.1).

Modified definitions from MHS Staff, Emotional Quotient Inventory 2.0 Manual (2011), Multi-Health Systems: Toronto. All rights reserved, Reproduced with permission of Multi-Health Systems, www.mhs.com.

The Four Branches of Emotional Intelligence

Mayer and Salovey proposed a model that identified four different factors of emotional intelligence: the perception of emotion, the ability to reason using emotions, the ability to understand emotion, and the ability to manage emotions.

1. *Perceiving emotions*: The first step in understanding emotions is to accurately perceive them. In many cases, this might involve understanding nonverbal signals such as body language and facial expressions. The ability to detect and decipher emotions in faces, pictures, voices, and cultural artifacts—including the ability to identify one's own emotions—is included in this factor. Perceiving emotions

Table 24.1 The EQ-i 2.0 scales and what they assess

EQ-i 2.0 scales	The EI skill assessed by each scale
Self-perception	
Self-regard	*Ability to respect and accept one's strengths and weaknesses*
Self-actualization	*Ability to improve oneself and pursue meaningful objectives*
Emotional self-awareness	*Ability to be aware of and understand one's feelings and their impact*
Self-expression	
Emotional expression	*Ability to express one's feeling verbally and nonverbally*
Assertiveness	*Ability to express feelings, beliefs, and thoughts in a nondestructive way*
Independence	*Ability to be self-directed and free of emotional dependency on others*
Interpersonal	
Interpersonal Relationship	*Ability to develop and maintain mutually satisfying relationships*
Empathy	*Ability to recognize, understand, and appreciate the feelings of others*
Social responsibility	*Ability to contribute to society, one's social group, to the welfare of others*
Decision making	
Problem solving	*Ability to solve problems where emotions are involved using emotions*
Reality testing	*Ability to remain objective by seeing things as they really are*
Impulse control	*Ability to resist or delay an impulse, drive, or temptation to act*
Stress management	
Flexibility	*Ability to adapt one's feeling, thinking, and behavior to change*
Stress tolerance	*Ability to effectively cope with stressful or difficult situations*
Optimism	*Ability to be remain hopeful and resilient, despite setbacks*
Additional scale	
Happiness	*Ability to feel satisfied with oneself, others, and life in general*

represents a basic aspect of emotional intelligence, as it makes all other processing of emotional information possible.

2. *Facilitating thought*: The next step involves using emotions to promote thinking and cognitive activity. Emotions help prioritize what we pay attention and react to; we respond emotionally to things that garner our attention. The ability to harness emotions to facilitate various cognitive activities, such as thinking and problem solving, is included in this factor. The emotionally intelligent person can capitalize fully upon his or her changing moods in order to best fit the task at hand.
3. *Understanding emotions*: The emotions that we perceive can carry a wide variety of meanings. If someone is expressing angry emotions, the observer must interpret the cause of their anger and what it might mean. For example, if your boss is acting angry, it might mean that he is dissatisfied with your work, or it could be because he got a speeding ticket on his way to work that morning or that he has been fighting with his wife. The ability to comprehend emotion language and to appreciate complicated relationships among emotions is included in this factor. For example, understanding emotions encompasses the ability to be sensitive to slight variations between emotions and the ability to recognize and describe how emotions evolve over time.
4. *Managing emotions*: The ability to manage emotions effectively is a key part of emotional intelligence. Regulating emotions, responding appropriately, and responding to the emotions of others are all important aspects of emotional management. The ability to regulate emotions in both ourselves and in others is included in this factor. Therefore, the emotionally intelligent person can harness emotions, even negative ones, and manage them to achieve intended goals.

According to Salovey and Mayer, the four branches of their model are "arranged from more basic psychological processes to higher, more psychologically integrated processes." For example, the lowest level branch concerns the (relatively) simple abilities of perceiving and expressing emotion. In contrast, understanding and managing emotions are more complex tasks.

Other Theories and Models

Other theories, mostly taken from one or both of the leading models of emotional intelligence in the field, have been proposed. Goleman, the *New York Times* writer, proposed his own theory of emotional intelligence, sometimes referred to as a "mixed model." He proposes the following aspects of emotional intelligence:

1. Self-awareness – the ability to know one's emotions, strengths, weaknesses, drives, values, and goals and recognize their impact on others while using gut feelings to guide decisions
2. Self-regulation – involves controlling or redirecting one's disruptive emotions and impulses and adapting to changing circumstances
3. Social skill – managing relationships to move people in the desired direction
4. Empathy – considering other people's feelings especially when making decisions
5. Motivation – being driven to achieve for the sake of achievement

The model introduced by Goleman focuses on emotional intelligence as a wide array of competencies and skills that drive leadership performance. He includes a set of "emotional competencies" within each construct of emotional intelligence. Emotional competencies are not innate talents, but rather learned capabilities that must be worked on and can be developed to achieve outstanding performance. Goleman posits that individuals are born with a general emotional intelligence that determines their potential for learning emotional competencies. Goleman's model of EI has been criticized in the research literature as mere "pop psychology" (Mayer et al. 2008).

In a 2010 meta-analysis, Dana Joseph and Daniel Newman proposed a new theory of emotional intelligence, called the "Cascading Model." It was originally created to address the prediction of job performance ratings. The model borrows concepts from the "Basic Models of Emotion" paradigm (Gross and Thompson 2007), as well as from Mayer and Salovey (1997). The model is also built on ability-based measures of emotional intelligence (as opposed to self-report measures). The authors offer criticisms of self-report measures (to justify using only ability-based measures), claiming that self-report measures lack scientific rigor and treat the term emotional intelligence as "…an umbrella term for a broad array of constructs that are connected only by their nonredundancy with cognitive intelligence…" (Joseph and Newman 2010a).

The new theory proposes that emotion perception, emotion understanding, and emotion regulation fit a progressive structure (where perception precedes understanding and understanding leads to conscious regulation). The perception of emotion includes perceiving the emotions of others and perceiving self-emotion. These two abilities may overlap greatly (Joseph and Newman 2010b). The understanding of emotion allows mediation between emotion perception and emotion regulation. However, it only applies to emotions which are consciously regulated, not emotions which are processed automatically (such as fear). The model also includes cognitive ability and both conscientiousness and emotional stability (from the Big Five personality traits) as factors which affect performance. Emotion perception, understanding, and regulation are all partial mediators for cognitive ability, conscientiousness, and emotional stability.

Measuring Emotional Intelligence

Methods of Measuring Emotional Intelligence

Generally, there are three methods used for measuring emotional intelligence.

The first is self-report inventories. Well-constructed self-report inventories compare an individual's responses to a database of thousands of others, preferably matched for age and gender, stratified based on census data, and include items reflecting how one sees oneself dealing with various situations, interacting with others, managing stress, perceiving the future, and using emotions in various ways—such as dealing with stress and making decisions. While many self-report inventories have appeared following the popularization of the concept, few have met the standards of test development as prescribed by the American Psychological Association (American Psychological Association 1999).

The second is 360° assessments. These add to the self-report inventory parallel report forms completed by others, such as one's work supervisor, subordinates, peers, clients, and even spouse. They all report on how they view the target individual in the same domains (e.g., emotional self-awareness, empathy, interpersonal skills, etc.). The idea behind 360 reports is that others are more likely to be candid about the target individual's true emotional skills. On the other hand, it has been argued that others may not intimately know what really goes on in the target's mind.

These methods are sometimes referred to as "trait-based approaches" or "mixed models." The trait-based/mixed model approach conceptualizes emotional intelligence as a set of emotion-related dispositions, attitudes, and self-perceptions located at the lower levels of the hierarchical personality taxonomy (Petrides and Furnham 2001; Petrides et al. 2007b). Like other personality variables, trait emotional intelligence is measured through self-report questionnaires, where respondents are asked to report on their typical beliefs, feelings, and behaviors. The most influential trait-based theory of EI has been Bar-On's (1997) model of emotional–social intelligence, operationalized with the Emotional Quotient Inventory (EQ-i; Bar-On 1997; EQ-i 2.0, MHS Staff 2011). The EQ-i assessed individual differences on a variety of traits and self-concepts organized into four broad EI dimensions: intrapersonal, interpersonal, adaptability, and stress management. The revised EQ-i 2.0 is organized into self-perception, self-expression, interpersonal, decision making, and stress management. Similar traits are measured by the Trait Emotional Intelligence Questionnaire (TEIQue) (Petrides and Furnham 2001; Petrides et al. 2007a) under the broad factors of emotionality, self-control, sociability, and well-being. However, this model lacks sufficient data for general use.

The third method of assessing emotional intelligence is through performance or ability measures. These instruments are designed much like traditional cognitive intelligence tests. They do not rely on one's opinions, thoughts, or feelings about oneself (or someone else's thoughts, feelings, or opinions) but rather on performance on well-normed tasks with known levels of difficulty. This method is known as an "ability-based approach."

According to John Mayer, "In regard to measuring emotional intelligence – I am a great believer that criterion-report (that is, ability testing) is the only adequate method to employ. Intelligence is an ability, and is directly measured only by having people answer questions and evaluating the correctness of those answers."

In their 2010 meta-analysis, Joseph and Newman (2010) attempted to build their theory on ability-based measures of emotional intelligence. In their conclusion, the authors ultimately concede that self-report models of emotional intelligence are stronger indicators of performance (such as job performance) than ability-based measures. This admission is supported by Bachman et al. (2000), who used a self-report measure (the Bar-On EQ-i) to compare more successful debt collectors to their less successful peers. They concluded that specific levels of emotional intelligence lead to enhanced job performance. However, in a large meta-analysis of the emotional intelligence literature, O'Boyle et al. (2010) suggest that each of the three methods are (a) significantly and positively correlated with job performance as a collective measure, (b) significantly and positively correlated with job performance individually, (c) not significantly different from each other in their ability to independently predict job performance, (d) correlated positively with extroversion, openness, agreeableness, conscientiousness, and cognitive ability, and (e) negatively related to neuroticism.

Table 24.2 Fifteen factors of the original EQ-i

Intrapersonal components	*Interpersonal components*
Self-regard (SR)	Empathy (EM)
Emotional self-awareness (ES)	Social responsibility (RE)
Assertiveness (AS)	Interpersonal relationship (IR)
Independence (IN)	
Self-actualization (SA)	
Adaptability components	*Stress management components*
Reality testing (RT)	Stress tolerance (ST)
Flexibility (FL)	Impulse control (IC)
Problem solving (PS)	
General mood components	
Optimism (OP)	
Happiness (HA)	

Scales for EI Assessment

Bar-On's EQ-i. A self-report test designed to measure competencies including awareness, stress tolerance, problem solving, and happiness. According to Bar-On, "Emotional intelligence is an array of noncognitive capabilities, competencies, and skills that influence one's ability to succeed in coping with environmental demands and pressures." Of the 15 factors underlying the EQ-i, 11 factors had been at the forefront of Bar-On's research for nearly two decades. The 15 factors of the original EQ-i (Bar-On 1997) and their abbreviations are in Table 24.2.

The EQ-i has been used extensively to successfully assess the effect of emotional intelligence on various measures of performance and quality of life.

Emotional intelligence has often been linked to employee selection, job satisfaction, financial performance, and sales performance (Porras and Anderson 1981; Spencer and Spencer 1993; Pesuric and Byham 1996; McClelland 1999; McEnrue and Groves 2006). Some newer research applies this knowledge to niche areas. Brown (2011) examined the relationship between nonprofit human service organization intellectual and financial capital and CEO emotional intelligence (measured using the EQ-i). Intellectual capital was correlated with self-regard and optimism, while total program effectiveness (a component of intellectual capital) was correlated with total EI score, an intrapersonal composite (consisting of self-regard, self-actualization, emotional self-awareness, assertiveness, and independence), as well as with flexibility, self-regard, and independence. Independence and reality testing were also found to be correlated with structural capital. Independence was also found to be related to total financial capital.

Enhelder (2011) also found a positive relationship (via correlation and linear regression) between total EI score and sales performance of financial advisors. They also found a correlation between sales performance and the assertiveness, independence, self-actualization, interpersonal relationships, optimism, and stress tolerance measures of the EQ-i. Additionally, in a multiple regression of the previously mentioned subscales, it was found that self-actualization predicts variability in sales performance.

Johnson (2011) investigated occupational well-being in law enforcement. Using the EQ-360, Johnson found that differences between managerial self-perceptions and employee perceptions of the manager's emotional intelligence were predictive of the employee's occupational well-being. The author further suggests that EI training can be used to encourage positive interaction between managers and employees generally, as well as between law enforcement personnel and the population they serve. Roth (2011) examined the connection between level of emotional intelligence in a pastor (EQ-i) and their ability to draw more attendees to their church (measured by increasing or decreasing attendance rates), finding that that five competencies were the best predictors of a pastor's success (emotional self-awareness, independence, flexibility, assertiveness and optimism).

Farnham (2012) found a positive relationship between emotional intelligence and sales performance in hospice sales professionals, finding that a 1-point increase in EQ-i total score leads to 0.316-unit increase in sales. Gender and tenure possessed no explanatory power in this study. Additionally, Mendelson

(2012) found that elite-level athletes (who are currently in the workplace) possess higher levels of emotional intelligence than their nonathlete coworkers and that number of years of experience is a predicting factor of the athlete's level of emotional intelligence.

The measure of EI (via the EQ-i) has also been involved in studies of education and academic performance. When comparing a high school student's EQ-i: Youth Version score to their GPA for that year, it was found that academic success was strong associated with the total EI score, the intrapersonal composite, the interpersonal composite, the adaptability, and the stress management (Parker et al. 2004a). These results are consistent with the findings of Parker et al. (2004b), who, using the EQ-i Short Form, found that these same EI dimensions (except for interpersonal abilities) were predictive of the successful transition from high school to first year university (measured by each student's academic performance in the first year).

In a study examining emotional intelligence (EQ-i), the teacher's sense of humor (Multidimensional Sense of Humor Scale), and each student's pretest versus posttest improvement on a standardized reading test, it was found that both the total EI and humor score of the teacher were related to improvements in academic performance, though EI held a greater explanatory power. Total EI was also related to the teacher's humor score (Fernandez 2011).

A study examining the relationship between EI (using the EQ-i Short Form) and the undergraduate clinical teaching effectiveness in a nursing faculty (using the Nursing Clinical Teacher Effectiveness Inventory) found that there is a significant positive relationship between total EI score and total teaching effectiveness score, as well as between many subscales of the two tools (Allen et al. 2012).

The EQ-i was also used to determine that the level of a teacher's emotional intelligence contributed to determining their students' success in a standardized test of mathematics (Shank 2012). The Urdu EQ-i: Youth Version found a positive correlation between academic achievement and emotional intelligence. Additionally, children from public schools were found to have a higher rating of EI than those in private schools (despite having an overall lower level of academic performance) (Malik and Shujja 2013).

Emotional intelligence has also been implicated as a factor in leadership abilities.

Jones (2012) examined the relationship between emotional intelligence (measured using the EQ-i) and leadership effectiveness (measured using the Leadership Practices Inventory) in senior level university sponsored administration professionals. Results indicated a significant correlation between EI and leadership effectiveness practices. Additionally, the "Enabling Others to Act" component of the Leadership Practices Inventory was correlated with the total EI score, the interpersonal composite score, empathy, social responsibility, impulse control, adaptability, problem solving, general mood, and optimism.

Osborne (2012) found a similar result in physicians, suggesting that physicians with higher emotional intelligence (EQ-i Short Form) showed better leadership practices (Leadership Practices Inventory). In regard to the "Enable Others to Act" component of the Leadership Practices Inventory, Osborne found that only the interpersonal composite was significantly correlated in physicians. There were no differences between males and females in these results.

MSCEIT. The Mayer-Salovey-Caruso Emotional Intelligence Test is an ability-based scale which measures how well individuals perform tasks (related to emotions) and solve emotional problems. The goal of the scale is to gather responses which represent the ability to solve emotional problems. Scores are unaffected by many confounding factors, such as self-concept, response set, and emotional state.

This scale is roughly based on the Multifactor Emotional Intelligence Scale (MEIS; Mayer et al. 2000). The MEIS included 402 items and four subscales: perception, assimilation, understanding, and managing emotion. The test itself was much like an IQ test, including possessing correct answers. The test also possessed two forms of scoring, including consensus scoring (which assesses the respondent's choice against the majority opinion based on hundreds of observations) and expert scoring (which compares the respondent's choice against the majority opinion of researchers who are experts in emotion-based research).

The MSCEIT provides 15 scores in total. The total emotional IQ score is an overall index of the test taker's overall emotional intelligence. Two "area scores" are also provided, consisting of an "experiential emotional intelligence score" (based on the respondent's ability to perceive emotional information, relate it to other sensations and use it to facilitate thought) and a "strategic emotional intelligence score" (based on the respondent's ability to understand emotional information and use it for planning and self-management purposes). There are four "branch scores," which correspond to the elements of the "Four Branches of Emotional Intelligence" theory. Finally, there are eight "task scores," which correspond to each of the either tasks found in the MSCEIT.

The MSCEIT has a strong reliability rating (Mayer, Salovey, Caruso, and Sitarenios, 2003). Bracket and Mayer (2001) found a test–retest reliability of $r = 0.86$ for the full-scale MSCEIT. Split-half estimates of the original test show a reliability of 0.91 for the full-scale test, 0.90 for the experiential area score, and 0.85 for the strategic area score. The MSCEIT also possesses strong validity ratings. Interrater reliability for face validity was $r = 0.83$. It also possesses strong structural validity (Mayer et al. 2001) and discriminant validity (Daus 2006).

Emotional Competence Inventory (ECI). Based on an older instrument known as the Self-Assessment Questionnaire, the ECI involves having people who know the individual offer ratings of that person's abilities on a number of different emotional competencies. The test contains 72 items, 4 dimensions (self-awareness, social awareness, self-management, and social skills), and 18 competencies (present only in ECI-2). It also features 360° assessment techniques, involving assessments of the individual from their peers and supervisor. Internal consistency for self-ratings ranges from 0.45 to 0.77, while the internal consistency for ratings from others ranges from 0.54 to 0.90 (Boyatzis and Sala 2004).

While many have used the ECI or ECI-2.0 in their work, there have been very few peer-reviewed articles published on the reliability and validity of the study (Cherniss 2000; Spencer 2001; Weinburger 2002; Conte and Dean 2006). The majority of these examinations are not favorable towards the ECI. The developers of the ECI suggest that validity evidence from the Self-Assessment Questionnaire acts as the evidence for the ECI's validity, though this is met with the criticism that these tests no longer hold merit, as the scale has been revised numerous times. For example, alpha coefficients reported for the first version of the ECI (Boyatzis and Burckle 1999) will not apply to the ECI-U (which has a different underlying structure) (Lewis et al. 2005). Those who have managed to examine the content of the ECI have noticed an overlap with four of the Big Five personality dimensions (conscientiousness, emotional stability, extraversion, and openness), as well as with other psychological concepts, such as self-awareness and self-confidence (Conte and Dean 2006). This has led to the suggestion that "…the ECI does not deserve serious consideration until peer-reviewed empirical studies using this measure are conducted" (Conte and Dean 2006). Since this 2006 quote, peer-reviewed papers have been published, though it provides (at best) modest support towards the validity of the ECI.

Byrne et al. (2007) argue that, while the ECI is moderately correlated with factors in the Big Five personality dimensions ($r = 0.22$–0.57), confirmatory factor analysis suggests that the factor structure of the ECI is different from that of the Big Five personality dimensions (which provides some evidence for discriminant validity for the ECI self-ratings). They also provide minimal evidence for convergent validity, by finding significant correlations ($r = 0.17$–0.25) between ECI self-ratings and the judges' ratings of emotional competency behaviors displayed by the individual during a "leaderless group discussion." However, due to the magnitude of the correlations found, they admit that the evidence is weak at best. Byrne and colleagues presented small correlations (0.11–0.29) between ECI scores and measures of work-related ratings (leaderless group discussion peer nominations, age-adjusted promotions, and coworker ratings of managerial skills), though they admit that ECI self-ratings predicted only one of the three criteria (coworker ratings of managerial skills) after controlling for personality and age. Finally, they find that there is no relationship between ECI self-ratings and general mental abilities, while other tests hold a relationship with general mental abilities, such as MSCEIT (Van Rooy and Viswesvaran 2004).

Lewis et al. (2005) used a confirmatory factor analysis to validate the ECI-U and determine whether four factors (self-awareness, self-management, social awareness, and relationship management) were identifiable in their medical student sample. They chose these four factors as Boyatzis and Goleman (2001) reported that the ECI-U to test 21 competencies as part of four clusters (consisting of these four factors). Statistical analyses testing this four-factor model suggested that the fit of this model to the data was very poor. Additionally, the internal consistencies of three of the four factors (measure by Cronbach's alpha) did not reach the minimum threshold for reliability. Also, there were large correlations between factors ($r = 0.999$ between self-awareness and self-management, and $r = 0.973$ between social awareness and relationship management), suggesting that the ECI-U's clusters were not entirely distinct from each other. The authors of this paper conclude that "… what this exercise has shown us is that we cannot be certain that we are measuring what this scale purports to measure i.e. EI" (p. 347).

Matthews et al. (2003) noted that ECI provides some utility, as it assesses many dissimilar concepts simultaneously. However, they additionally state that there are more advanced techniques for assessing many of the factors found in the ECI. Leslie and Fleenor (1998) also note that other 360° rating instruments exist, which measure similar skills to the ECI. Proponents of the ECI have yet to show that the ECI 360° ratings are different from other well-known 360° rating instruments. Finally, there have been comparisons between the Bar-On EQ-i (a self-report measure) and ability-based measures (such as the MSCEIT) (Mayer et al. 2000; Brackett and Mayer 2003). The ECI literature lacks these comparisons.

McEnrue and Groves (2006) criticize the lack of validity in the ECI-2. The authors question the content validity of the test, claiming that several competencies which Goleman identified in his theory seem to be products (not dimensions) of emotional intelligence. They also state that the ECI-2 seems to assess competencies which are not emotional (such as serving as a catalyst for change) and does not assess competencies which are related to emotionally intelligent individuals (such as knowledge of when to express emotions). McEnrue and Groves also question the construct validity of the ECI-2, stating that the measured competencies are significantly related to many existing personality indexes (thus, the test offers little value beyond already existing tests). They also claim that face validity is also only supported by anecdotal information. The authors criticize the external validity of the ECI-2, drawing attention to the lack of normative data and compiled demographic information of respondents. Additionally, the ECI-1 and ECI-2 are very different from each other, thus supporting evidence for the external validity of the ECI-1 cannot apply to the ECI-2.

Finally, unlike the EQ-i, the ECI lacks findings which allow the prediction of performance and satisfaction in various areas of life. The ECI was found to have significant overlap in most of its dimensions and contributed only a small amount of variance above the NEO-PI in predicting organizational performance (Murensky 2000). Results of the ECI were compared before and after training school teachers with a program designed to increase awareness and understanding of cognitive messages which create feelings and behaviors. There were no significant differences pre- and posttest, for both the experimental and control group (Walker 2001). In a study involving emotional intelligence and an individual's tolerance levels towards workplace bullying, it was found that none of the four competency clusters nor the total score from the ECI was related to an individual's tolerance (Roundy 2007).

In a study examining the connection between an ESCI score and leadership style (via the Multifactor Leadership Questionnaire), no significant relationship was found. However, it was found that those who score higher on the ESCI *believe* that they practice a form of leadership closer to a transformational leadership style (as compared to the alternatives of transactional leadership and passive avoidance) (Woods 2010). Self-awareness factors of the ECI-2.0 do not predict the scores on a Burnout scale (Maslach Burnout Inventory—GS) in chefs, while the conscientiousness factors from the NEO were able to predict the Burnout scale score (Hintertoisser 2011). Finally, only the influence competency component of the ECI is significantly related to general job satisfaction (Minnesota Satisfaction Questionnaire), leading the authors to conclude that there is a lack of sufficient evidence to conclude a meaningful relationship between emotional intelligence competencies (as tested by the ECI) and general job satisfaction (Agbolou 2011).

EQ 360 2.0. The EQ 360 2.0 is a 360° assessment (also known as a multirater or multisource feedback assessment), based on the EQ-i 2.0, which provides the individual with feedback from their peers and supervisors, allowing them a deeper understanding of their performance from multiple perspectives. This method of assessment provides the means for comparing internal perceptions with external perceptions and acts as a means for enhanced self-awareness and subsequent behavioral change.

The normative group consists of 3,200 ratees (59.2 % female). Ratees are demographically similar to the 2008 US and 2006 Canadian census results. Data from both countries are included in the normative sample, as there were no statistically significant differences between the ratings provided by the US and Canadian participants.

The EQ 360 2.0 possesses high reliability ratings. In regard to the internal consistency of the normative sample, the Cronbach's alpha is at least 0.85 for all but two subscales. The exceptions are emotional expression ($\alpha = 0.82$) and assertiveness ($\alpha = 0.79$). This suggests that the items cohesively measure total EI, as well as the construct measured by each of the scales and subscales. Test–retest reliability was calculated ($n = 203$), producing correlations for total and subscale scores ranging from $r = 0.76$ to 0.89. Test–retest stability values were also calculated, with the results suggesting that 90 % or more of the individual's scores (for both the total score and each subscale score) did not change by more than one standard deviation between the two testing periods.

The EQ 360 2.0 also possesses high validity ratings, as determined by multiple tests. The first set of tests examined the correlations between the composite scales and the subscales within the normative sample, in an attempt to determine if the pattern of results found in the EQ-i 2.0 would be replicated. The resulting correlations ranged from $r = 0.64$ (self-expression/interpersonal) to 0.86 (decision making/stress management). In most cases, the correlations were stronger than the corresponding correlation for the EQ-i 2.0 normative sample. Additionally, subscales within the same composite scale possessed strong correlations. Every value exceeded a medium effect size, and most exceeded a large effect size. Correlations ranged from $r = 0.37$ (emotional express/independence) to 0.81 (empathy/interpersonal relationships). These results suggest that the composite scales and subscales share a relevant underling factor. Also, these results are similar to the results of this analysis for the EQ-i 2.0.

The second set of tests examined the associations between 108 pairs of self-ratings and the ratings of their peer (often family members, spouses, or friends who interacted with the participants frequently and knew the participant very well). The correlation between the total scores was $r = 0.60$ ($p < 0.01$) with correlations ranging from $r = 0.44$ (stress tolerance) to 0.72 (happiness). Correlations for the composite scales and subscales were all significant at $p < 0.01$. Almost every correlation reached the criterion for a large effect size. The observed pattern suggests that there is strong agreement between self-report and the report of others and that EI (as measured by the EQ-i 2.0 and the EQ 360 2.0) is a robust trait that is evaluated similarly via both self-report and report from external observers. However, these correlations are not high enough to suggest redundancy between the two forms of report (especially as self-reports will not always align with external observer ratings). Each measure provides unique and important information about the individual.

In order to supplement the above correlational results, the standard scores of the EQ-i 2.0 and the EQ 360 2.0 were compared. By subtracting the EQ 360 2.0 standard score from the EQ-i 2.0 score, a difference score was determined. It was previously established that a difference of 10 points would result in a meaningful difference between the two standard scores (Multi-Health Systems 2011, see *EQ-i 2.0 User's Handbook*, Chapter 6). Over half of the difference scores (for total EI score, composite scales, and subscales) had a value of less than 10, suggesting a good degree of consistency between the EQ-i 2.0 and EQ 360 2.0 scores. However, given that almost half of the difference scores were meaningful, both scores are still important to collect in order to build a complete assessment of the individual.

The third set of tests examined the relationship between emotional intelligence and general/social adjustment. Social adjustment was measured by the Social Adjustment Scale – Self-Report (SAS-R from Weissman 1999) and compared to the EQ-i 2.0 and EQ 360 2.0. Correlations between the SAS-R and the EQ-i 2.0/EQ 360 2.0 were mostly strong. Stepwise multiple regression revealed that the SAS-R total score is independently related to the total EI score for both the EQ-i 2.0 and the EQ 360 2.0, as well as most of the composite scales and subscales ($p < 0.05$). In other words, the ratings in the self-report data and the external observer data are both uniquely informative about the individual's SAS-R total score. Additionally, by examining the R2 change values in the stepwise regression, it can be observed that each of the subscales and composite scales from the EQ-i 2.0 and the EQ 360 2.0 provided unique and incremental contributions towards social adjustment.

Finally, the potential of an ethnicity bias effect in the rater–ratee relationship was examined. Ideally, ethnicity should not affect the score for either the rater or the ratee. Analysis of covariance techniques were used to examine potential effects on the EQ 360 2.0 total score, using rater and ratee ethnicity as independent variables, as well as ratee gender and age group as covariates. Multivariate analyses of covariance were also used to examine the effects of the above independent variables and covariates on the composite scales and subscales of the EQ 360 2.0. None of the analyses produced significant results, suggesting that the raters did not show differences in their ratings based on the ethnicity of the ratee. Additionally, the Wilks' Lambda values suggested that a negligible amount of variance could be explained by the interaction between rater and ratee ethnicity, while none of the effect sizes met the requirements to be classified as a small effect size.

Trait Meta-Mood Scale (TMMS). The Trait Meta-Mood Scale (Salovey et al. 1995) was the very first self-report measure to be applied in EI-related research. The 30-item scale is based on the earlier work on the cognitions accompanying subjective mood experiences (Mayer and Gaschke 1988) and measures three core meta-mood processes: attention (attending to and monitoring changes in one's mood states), clarity (discerning the nature and causes of one's feelings), and repair (regulating one's emotions in adaptive ways). The TMMS possesses sound reliability (Cronbach's $\alpha = 0.70$) and good validity, supported by a sizeable body of evidence supporting the three-factor structure. It also possesses strong predictive utility for a range of performance and health-related outcomes (e.g., Gignac et al. 2003; Salovey et al. 2002; Thompson et al. 2007). Its main shortcoming, however, is that the TMMS was never designed to measure the EI construct specifically (Salovey et al. 1995). As a result, several components of major EI models (such as empathy, understanding emotions of others, interpersonal skills) are overlooked, which limits the scope of its utility in theoretical and applied research (Parker et al. 2011).

Assessing Emotions Scale (AES). This is another brief self-report measure that has been widely used in EI-related research (Schutte et al. 1998). The AES is based on Salovey and Mayer's (1990) original definition of EI as comprising four ability domains (appraising emotions in self and others, expressing emotions, managing emotions of self and others, and utilizing emotions in problem solving). This scale has become the choice for many applied EI researchers, due to its solid theoretical foundation and the concise nature of the scale (Schutte et al. 2009). However, the AES has several psychometric problems, including an unreliable measurement structure. In some factor analytic investigations (Brackett and Mayer 2003; Schutte et al. 1998), AES responses were found to be represented adequately by a single common factor, whereas in others (Keele and Bell 2008; Petrides and Furnham 2000; Saklofske et al. 2003) four separate factors were identified, although their item composition varied from study to study and did not fully correspond to the domains in Salovey and Mayer's conceptual model (Parker et al. 2011). Given these issues, most applications use only the total score, as this measure is a reliable (Cronbach's $\alpha = 0.80$) but nonspecific index of global EI (Schutte et al. 2009).

EQ-i and TEIQue–Short Forms. The 35-item Emotional Quotient Inventory–Short Form (EQ-i:S; Bar-On 2002) and the 30-item Trait Emotional Intelligence Questionnaire–Short Form (TEIQue–SF; Petrides and Furnham 2006) are short-form alternatives to their respective full length assessments. Neither of the two short forms has been researched as extensively as the TMMS or the AES, their well-known theoretical bases make them worthwhile measures to consider. Both short forms were designed to match the higher-order measurement structure of their respective parent scales, with items selected for inclusion on the basis of item-total correlations and representative coverage of the facet level content (Parker et al. 2011). The EQ-i:S yields a total EI score and four subscale scores (intrapersonal, interpersonal, adaptability, and stress management), and the TEIQue–SF provides a total trait EI score and four-factor scores (emotionality, self-control, sociability, and well-being).

The four-factor structure of the EQ-i:S scores has been confirmed in a large sample of adults (Bar-On 2002), with each subscale possessing internal consistency (Cronbach's $\alpha = 0.70$) and congruence with scores on the corresponding long-form scale ($r = 0.73$–0.93). Parker et al. (2011) provided evidence that EQ-i:S scores are more strongly associated with measures of conceptually similar constructs (such as the MSCEIT and the 20-item Toronto Alexithymia Scale) than they are with more general dimensions of personality. Additionally, at the latent-variable level, the EQ-i:S also shares more conceptual overlap with the MSCEIT (65 % shared variance) than alexithymia (41 % shared variance). They concluded that "…the EQ-i:S produced internally consistent, temporally reliable and theoretically meaningful responses that also followed a stable, gender-invariant, multidimensional measurement structure" (p. 773). On the other hand, the TEIQue–SF shows low levels of internal consistency at the individual factor level (Cronbach's $\alpha = 0.70$). The developers of the scale have recommend the use of the total scale score (Cronbach's $\alpha = 0.80$) instead of the individual factors (Petrides 2009; Petrides and Furnham 2006). Like the AES, the TEIQue–SF is useful only as a general measure of global EI, while the EQ-i:S offers a unique advantage over other brief EI scales and can be used for a wider range of research questions and applications (Parker et al. 2011).

Neurological Correlates of Emotional Intelligence

A wealth of information has been generated over recent years concerning the emotional implications for performance success and the role emotions play in survival and everyday decision making. Recent technological advances in the study of emotion (e.g., fMRI, EEG, etc.) have shed light on the cortical and subcortical structures of the brain linked to the emotion network that drives how we think, feel, and act. Of particular interest are the subcortical structures of the thalamus, cingulate cortex, amygdala, and the cortical regions, such as the temporal lobe and the prefrontal cortex. In the event one of these areas should become impaired, so too does the individual's ability to process or use information effectively, often affecting thoughts, feelings, and actions. This structure–function relationship maintains implications for stress tolerance, well-being, decision making, and success. Recently, some researchers have attempted to link EI to the very neural mechanisms that have been demonstrated to have a direct impact on what we perceive, how we react, the decisions we make, and ultimately the quality of life we lead.

According to the somatic marker hypothesis (Damasio et al. 1991), when we make a decision, we first weigh the pros and cons or the benefits/consequences of the response options, a process that requires both emotional and cognitive processing. However, when faced with a simple choice, we often resort to cognitive rules to assist in the decision process; as the decision becomes more complex or we place greater value on the outcome, our cognitive rules may not be sufficient to render a decision and we get stuck. Enter emotions and the somatic marker hypothesis.

Somatic markers are simply connections between a stimulus (i.e., choice) and a resulting physiological sensation we experience when presented with such a stimulus. When presented with a given stimulus, we experience certain sensations which in turn bias or influence our decisions. In most cases, the somatic marker directs attention to the most meaningful information to enhance decision making (Damasio et al. 1991); however, in the event a deficit in emotional processing is evident, decision making and judgment become impaired.

Bar-On et al. (2003) set out to determine whether individuals with impaired somatic markers (i.e., lesions to the ventromedial prefrontal cortex; vmPFC) would reflect that impairment through abnormal emotional intelligence. In this case, damage to the vmPFC is often associated with impaired judgment and decision making and, in turn, should be reflected by lower scores in EI. Comparing patients with lesions to various brain regions, Bar-On et al. (2003) demonstrated that those with lesions to the vmPFC reported lower EI despite showing no difference in IQ. Given the implications of vmPFC in the somatic marker hypothesis for decision making and subsequent behavior, Killgore and Yurgelun-Todd (2007) set out to assess the link between vmPFC activation (measured using fMRI) and levels of emotional intelligence. Similar to the results of Bar-On et al. (2003), Killgore and Yurgelun-Todd reported that adolescents with relatively low emotional intelligence respond to emotionally provocative pictures with greater and more extensive brain activation than do those with well-developed emotional intelligence. In other words, emotional intelligence can moderate the impact of stressful stimuli, allowing the brain to operate more efficiently under stressful conditions (Haier et al. 1992).

Damage to temporal lobe functioning has been reliably linked to increased agitation, difficulty managing emotions, heightened irritability, and, more recently, impaired social cognition (Walpole et al. 2008). If damage or functional impairment of the temporal lobe presents emotional challenges, then perhaps the temporal region of the brain is linked to EI. Walpole et al. (2008) in a controlled experiment measured the emotional intelligence of patients with temporal lobe epilepsy and a healthy cohort matched for age and IQ. These findings suggest that impairment to the medial temporal lobe is related to lower emotional intelligence and impaired facial recognition and to greater psychological distress as compared to healthy individuals.

To this point, emotional intelligence has been linked to vmPFC and the temporal lobe by means of comparing relatively healthy individuals to those with a structural deficit. In a series of studies conducted by Killgore and colleagues (Kahn-Greene et al. 2006; Killgore et al. 2007, 2008), healthy participants were used in a repeated measures design to isolate the simulated effects of vmPFC impairment on emotion functioning. Killgore et al. (2008) were able to demonstrate the cortical connection of emotional intelligence to the prefrontal cortex via sleep deprivation. It has been shown that sleep deprivation can result in temporary impairment of the prefrontal cortex, resulting in difficulties regulating higher-order executive functions such as impulse control, inhibition of aggression, willingness to act in a socially acceptable way (Kahn-Greene et al. 2006), and moral judgment (Killgore et al. 2007). When comparing sleep-deprived EQ-i results to baseline results, sleep-deprived participants reported decreased total EQ, intrapersonal, interpersonal, and stress management composite scores (Killgore et al. 2008), and those scoring lower at baseline were more susceptible to decrements in moral judgment performance (Killgore et al. 2007).

More recently, regional brain volumes have been linked to emotional intelligence. Killgore and colleagues (2012) examined the correlation between gray matter volume in the somatic marker circuitry (Bar-On et al. 2003) and the ability-based model of emotional intelligence (via the MSCEIT), as well as a self-report measure of emotional intelligence (the Bar-On EQ-i). Their results suggest that the complete ability-based model of emotional intelligence is positively correlated with the volume of gray matter in the left posterior insula, while the complete self-report model was not correlated with gray matter volume within the somatic market circuitry. Additionally, individual components of each model correlated with gray matter volume. The strategic emotional intelligence subscale of the MSCEIT was correlated with the gray matter volume of the bilateral medial prefrontal cortex, the left posterior and anterior insula, and the ventrolateral prefrontal cortex area (including the inferior frontal gyrus). The stress management subscale of the EQ-i was positively correlated with the gray matter volume of both the right and left vmPFC. It is interesting to note that the volume of the amygdala is not related to either the ability or self-report models of emotional intelligence, despite its importance within the somatic marker circuitry. However, the authors also note that the posterior insula, the anterior insula, and the vmPFC are also important to the somatic market circuitry and are positively correlated with the tasks used in this study. Specifically, the posterior insula influences somatic processing (Craig 2003) and emotional processing (Xue et al. 2010), the anterior insula allows the integration of cognition and emotion (Gu et al. 2012), and the vmPFC is associated with understanding and managing emotional information (via the MSCEIT strategic emotional intelligence correlation), as well as emotional control (via the EQ-i stress management correlation).

Aron Barbey and his colleagues (2012) investigated the idea that the neural architecture of emotional and social intelligence overlaps that of cognitive intelligence. Their sample consisted of 152 males who experienced penetrating head injuries (leading to brain lesions in a variety of locations) during the Vietnam War. Each participant was administered the MSCEIT as a measure of emotional intelligence, the WAIS-III (Wechsler 1997) as a measure of cognitive intelligence, and the NEO-PI-R (Costa and McCrae 1997) as a measure of personality traits.

Emotional intelligence deficits (lower scores on the MSCEIT) were associated with damage to the social cognitive network (review in Saxe 2006), suggesting that social and emotional processes are integrated at the neural level. Specific areas of damage included the left posterior temporal cortex (associated with recognition of the form of human bodies), the left posterior superior temporal sulcus (associated with interpreting the motions of the human body in relation to the individual's goals), the left temporoparietal junction (associated with the ability to reason about the contents of mental states), and the left orbitofrontal cortex (associated with the support of emotional empathy and triadic relations between two minds and an object, as well as supporting shared attention and collaborative goals). Additionally, there were damaged white matter fiber tracts which are associated with the social cognitive network, including the super longitudinal/arcuate fasciculus (which connects the temporal lobe, parietal lobe, and inferior area of the frontal lobe), the superior fronto-occipital fasciculus (which connects the dorsolateral prefrontal cortex and the frontal area of the superior parietal cortex), and the uncinate fasciculus (which connects the anterior temporal cortex and the amygdala with the orbitofrontal and frontopolar temporal cortex).

Emotional intelligence deficits were also associated with impaired behavioral performance in the verbal comprehension component of the WAIS-III, suggesting that emotional intelligence shares neural systems which are essential for cognitive or crystallized intelligence. Selective damage was observed in the left hemisphere perisylvian language network (review in Hickok and Poeppel 2007). Impaired performance was associated with both the ventral and the dorsal pathways of this network. The ventral pathway (consisting of the anterior and posterior middle temporal gyrus, as well as the middle posterior superior temporal sulcus) is known for its role in language comprehension, including the support of mapping sensory or phonological representations onto lexical or conceptual representations. The dorsal pathway (consisting of the anterior and posterior insula, as well as a section along the parietotemporal boundary) is known for its role in language production, including the support of converting sensory or phonological representations into motor articulations.

Emotional intelligence deficits were also found in relation to deficits in processing speed (via the WAIS-III). Deficits in both processing speed and emotional intelligence were related to damage to the left dorsolateral prefrontal cortex (which is involved in the regulation and control of both emotional and social behavior).

Deficits in emotional intelligence were found to be linked to only the conscientiousness personality trait of the Big Five personality traits. Brain regions associated with the impairments are involved in the regulation and control of behavior and are also implicated in social information processing networks (review in Saxe 2006). These regions include the right dorsolateral prefrontal cortex, the left orbitofrontal cortex, and the left temporoparietal junction.

Finally, when residual scores (which removed variance shared with significant cognitive and personality predictors) were analyzed, additional brain structures were shown to be involved in emotional intelligence. These areas showing additional activation include the right orbitofrontal cortex, the left inferior and superior parietal cortex, and the major white matter fiber tracts (including the superior longitudinal/arcuate fasciculus, the superior fronto-occipital fasciculus, and the uncinate fasciculus).

Overall, Barbey and colleagues found that the network involved in emotional intelligence was distributed in nature, though the core components consisted in either white matter areas of the brain (suggesting the importance of communication between the employed brain regions for greater emotional intelligence) or regions associated with social information processing (suggesting the importance in the regulation and control of social behavior in emotional intelligence). They also suggest that the orbitofrontal cortex is a central component of the network for emotional intelligence, especially as it has been suggested as an important region for emotional and social cognition (review in Kringelbach 2005). Researchers have proposed that the medial orbitofrontal cortex is crucial for emotional intelligence processes, as it is involved in monitoring emotional properties of social and environmental stimuli.

The results of recent neurological investigations show promise for enhancing our understanding of the role emotional intelligence plays in our capacity to meet our daily demands. Emotional intelligence appears to have a broad neurological representation which, in turn, serves to moderate the effects of emotional stimuli affecting both our decisions and behaviors.

References

Agbolou, K. (2011). *Relationship between emotional intelligence and job satisfaction: A correlational analysis of a retail organization* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3471686).

Allen, D. E., Ploeg, J., & Kaasalainen, S. (2012). The relationship between emotional intelligence and clinical teaching effectiveness in nursing faculty. *Journal of Professional Nursing, 28*(4), 231–240.

American Psychological Association. (1999). Standards for educational and psychological testing. Developed jointly by the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME), American Psychological Association, Washington, DC.

Bachman, J., Stein, S., Campbell, K., & Sitarenios, G. (2000). Emotional intelligence in the collection of debt. *International Journal of Selection and Assessment, 8*(3), 176–182.

Barbey, A. K., Colom, R., & Grafman, J. (2012). Distributed neural system for emotional intelligence revealed by lesion mapping. *Social Cognitive and Affective Neuroscience*, 9, 265–272

Bar-On, R. (1988). *The development of a concept of psychological well-being*. Unpublished doctoral dissertation, Rhodes University, South Africa.

Bar-On, R. (1997). The Bar-On Emotional Quotient Inventory (EQ-i): Technical manual. Toronto: Multi-Health Systems.

Bar-On, R. (2002). *Bar-On Emotional Quotient Inventory: Short (EQ-i:S): Technical manual*. Toronto: Multi-Health Systems.

Bar-On, R., & Parker, J. D. A. (2000). *The Bar-On Emotional Quotient Inventory: Youth Version (EQ-i:YV) Technical manual*. Toronto: Multi-Health Systems.

Bar-On, R., Tranel, D., Denburg, N. L., & Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain, 126*, 1790–1800.

Boyatzis, R. E., & Burckle, M. (1999). *Psychometric properties of the ECI*. Boston: Hay/McBer Group.

Boyatzis, R. E., & Goleman, D. (2001). Emotional competence inventory – University version. Self-assessment questionnaire, profile and interpretive notes. London: Hay Resources Direct.

Boyatzis, R. E., & Sala, F. (2004). Assessing emotional intelligence competencies. In G. Geher (Ed.), *The measurement of emotional intelligence: Common ground and controversy* (pp. 147–180). Hauppauge: Nova Science.

Brackett, M., & Mayer, J. D. (2001). Comparing measures of emotional intelligence. Paper presented at the Third Positive Psychology Summit, Washington, DC.

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29*, 1147–1158.

Brown, J. S. (2011). *Emotional intelligence exploration: A potential relationship between CEO EI and non-profit intelligence and financial capital* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3468350).

Byrne, J. C., Dominick, P. G., Smither, J. W., & Reilly, R. R. (2007). Examination of the discriminant, convergent and criterion-related validity of self-ratings on the emotional competence inventory. *International Journal of Selection and Assessment, 15*(3), 341–353.

Cherniss, C. (2000, April 15). *Emotional intelligence: What it is and why it matters.* Paper presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, New Orleans.

Conte, J. M., & Dean, M. A. (2006). Can emotional intelligence be measured? In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 59–77). Mahwah: Lawrence Erlbaum Associates.

Costa, P. T., Jr., & McCrae, R. R. (1997). Stability and change in personality assessment: The revised NEO personality inventory in the year 2000. *Journal of Personality Assessment, 68*, 86–94.

Craig, A. D. (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology, 13*(4), 500–505.

Damasio, A. R., Tranel, D., & Damasio, H. (1991). Somatic markers and the guidance of behaviour: Theory and preliminary testing. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), Frontal lobe function and dysfunction (pp. 217–229). New York: Oxford University Press.

Daus, C. S. (2006). The case for an ability-based model of emotional intelligence. In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 301–324). Mahwah: Lawrence Erlbaum Associates.

Enhelder, M. (2011). *Emotional intelligence and its relationship to financial advisor sales performance* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3465842).

Farnham, B. (2012). *Exploring impacts of emotional intelligence, gender and tenure on sales performance among hospice sales professionals* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3501884).

Fernandez, S. (2011). *The relationship between teachers' emotional intelligence and sense of humor, and student achievement* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3467199).

Gardner, H. (1976). *The shattered mind*. London: Routledge & Kegan Paul.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gignac, G. E., Palmer, B., Bates, T., & Stough, C. (2003). Differences in confirmatory factor analysis model close-fit index estimates obtained from AMOS 4.0 and AMOS 5.0 via full information maximum likelihood–no imputation: Corrections and extension to Palmer et al. (2003). *Australian Journal of Psychology, 58*(3), 144–150.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation*. New York: Guilford Press.

Gu, X., Liu, X., Van Dam, N. T., Hof, P. R., & Fan, J. (2012). Cognition-emotion integration in the anterior insular cortex. *Cerebral Cortex*, *23*(1), 20–27.

Haier, R. J., Siegel, B., Tang, C., Abel, L., & Buchsbaum, M. S. (1992). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence, 16*, 415–426.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*, 393–402.

Hinterstoisser, T. M. (2011). *An investigation into the relationship between conscientiousness, self-awareness and occupational stress outcomes in culinary chefs* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3478296).

Johnson, A. M. (2011). *The impact of managerial emotional intelligence perceptions on the occupational well-being of employees in a police department* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3482199).

Jones, V. D. (2012). *The relationship between emotional intelligence and leadership effectiveness among sponsored research administrators* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3514683).

Joseph, D. L., & Newman, D. A. (2010a). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*(1), 54–78.

Joseph, D. L., & Newman, D. A. (2010b). Discriminant validity of self-reported emotional intelligence: A multitrait–multisource study. *Educational and Psychological Measurement, 70*(4), 672–694.

Kahn-Greene, E. T., Lipizzi, E. L., Conrad, A. K., Kamimori, G. H., & Killgore, W. D. (2006). Sleep deprivation adversely affects interpersonal responses to frustration. *Personality and Individual Differences, 41*(8), 1433–1443.

Keele, S. M., & Bell, R. C. (2008). The factorial validity of emotional intelligence: An unresolved issue. *Personality and Individual Differences, 44*, 487–500.

Killgore, W. D., & Yurgelun-Todd, D. A. (2007). Neural correlates of emotional intelligence in adolescent children. *Cognitive, Affective, & Behavioral Neuroscience, 7*(2), 140–151.

Killgore, W. D., Killgore, D. B., Day, L. M., Li, C., Kamimori, G. H., & Balkin, T. J. (2007). The effects of 53 hours of sleep deprivation on moral judgment. *Sleep, 30*(3), 345–352.

Killgore, W. D., Kahn-Greene, E. T., Lipizzi, E. L., Newman, R. A., Kamimori, G. H., & Balkin, T. J. (2008). Sleep deprivation reduces perceived emotional intelligence and constructive thinking skills. *Sleep Medicine, 9*(5), 517–526.

Killgore, W. D., Weber, M., Schwab, Z. J., DelDonno, S. R., Kipman, M., Weiner, M. R., & Rauch, S. L. (2012). Gray matter correlates of trait and ability models of emotional intelligence. *NeuroReport, 23*(9), 551–555.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience, 6*, 691–702.

Landy, F. L. (2005). Some historical and scientific issues related to research on emotional intelligence. *Journal of Organizational Behaviour, 26*(4), 411–424.

Leeper, R. W. (1948). A motivational theory of emotions to replace "emotions as disorganized responses". *Psychological Review, 55*, 5–21.

Leslie, J. B., & Fleenor, J. W. (1998). *Feedback to managers: A review and comparison of multi-rater instruments for management development*. Greensboro: Center for Creative Leadership.

Lewis, N. J., Rees, C. E., Hudson, J. N., & Bleakley, A. (2005). Emotional intelligence in medical education: Measuring the unmeasurable? *Advances in Health Sciences Education, 10*, 339–355.

Malik, F., & Shujja, S. (2013). Emotional intelligence and academic achievement: Implications for children's performance in schools. *Journal of the Indian Academy of Applied Psychology, 39*(1), 51–59.

Matthews, G., Zeidner, M., & Roberts, R. (2003). *Emotional intelligence: Science and myth*. Cambridge, MA: The MIT Press.

Mayer, J. D. (1986). How mood influences cognition. In N. E. Sharkey (Ed.), *Advances in cognitive science* (pp. 290–314). Chichester: Ellis Horwood Limited.

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology, 55*, 102–111.

Mayer, J. D., & Salovey, P. (1988). Personality moderates the interaction of mood and cognition. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behavior* (pp. 87–99). Toronto: Hogrefe.

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence, 17*(4), 433–442.

Mayer, J. D., & Salovey, P. (1995). Emotional intelligence and the construction and regulation of feelings. *Applied and Preventive Psychology, 4*, 197–208.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators*. New York: Basic Books.

Mayer, J. D., & Volanth, A. J. (1985). Cognitive involvement in the mood response system. *Motivation and Emotion, 9*, 261–275.

Mayer, J. D., Mamberg, M. H., & Volanth, A. J. (1988). Cognitive domains of the mood system. *Journal of Personality, 56*, 453–486.

Mayer, J. D., DiPaolo, M. T., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment, 54*, 772–781.

Mayer, J. D., Salovey, P., Gomberg-Kaufman, S., & Blainey, K. (1991). A broader conception of mood experience. *Journal of Personality and Social Psychology, 60*, 100–111.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2000). Selecting a measure of emotional intelligence: The case for ability testing. In R. Bar-On & J. D. A. Parker (Eds.), *Handbook of emotional intelligence* (pp. 320–342). New York: Jossey-Bass.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion, 1*(3), 232–242.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 3*(1), 97–105.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry, 15*(3), 197–215.

Mayer, J. D., Barsade, S. G., & Roberts, R. D. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology, 59*, 507–536.

McClelland, J. L. (1999). Cognitive modeling: Connectionist. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 137–141). Cambridge, MA: MIT Press.

McEnrue, M. P., & Groves, K. (2006). Choosing among tests of emotional intelligence: What is the evidence? *Human Resource Development Quarterly, 17*(1), 9–42.

McGrath, R. E. (2005). Conceptual complexity and construct validity. *Journal of Personality Assessment, 85*, 112–124.

Mendelson, R. A. (2012). *From athlete to employee: A quantitative study of the impact of emotional intelligence in the workplace – Elite athletes versus nonathletes* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3524312).

Multi-Health Systems Inc. (2011). Emotional Quotient Inventory 2.0: User's handbook. Toronto: Multi-Health Systems.

Murensky, C. L. (2000). *The relationships between emotional intelligence, personality, critical thinking ability and organizational leadership performance at upper levels of management* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (9962991).

Murphy, K. R. (2006). *A critique of emotional intelligence: What are the problems and how can they be fixed?* Mahwah: Lawrence Erlbaum Associates.

O'Boyle, E. H., Jr., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2010). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior, 32*(5), 788–818.

Osborne, A. M. (2012). *The relationship between emotional intelligence and leadership practices among physicians* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3523050).

Parker, J. D. A., Creque, R. E., Sr., Barnhart, D. L., Harris, J. I., Majeski, S., Wood, L. M., Bond, B. J., & Hogan, M. J. (2004a). Academic achievement in high school: Does emotional intelligence matter? *Personality and Individual Differences, 37*, 1321–1330.

Parker, J. D. A., Summerfeldt, L. J., Hogan, M. J., & Majeski, S. (2004b). Emotional intelligence and academic success: Examining the transition from high school to university. *Personality and Individual Differences, 36*, 163–172.

Parker, J. D. A., Keefer, K. V., & Wood, L. M. (2011). Toward a brief multidimensional assessment of emotional intelligence: Psychometric properties of the emotional quotient inventory—Short form. *Psychological Assessment, 23*(3), 762–777.

Pesuric, A., & Byham, W. (1996). The new look in behavior modeling. *Training and Development, 50*(7), 24–30.

Petrides, K. V. (2009). *Technical manual for the Trait Emotional Intelligence Questionnaires (TEIQue)*. London: London Psychometric Laboratory.

Petrides, K. V., & Furnham, A. (2000). On the dimensional structure of emotional intelligence. *Personality and Individual Differences, 29*, 313–320.

Petrides, K. V., & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality, 15*, 425–448.

Petrides, K. V., & Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology, 36*, 552–569.

Petrides, K. V., Furnham, A., & Mavroveli, S. (2007a). Trait emotional intelligence: Moving forward in the field of EI. In G. Matthews, M. Zeidner, & R. Roberts (Eds.), Emotional intelligence: Knowns and unknowns (Series in affective science). Oxford: Oxford University Press.

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007b). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology, 98*, 273–289.

Porras, J. I., & Anderson, B. (1981). Improving managerial effectiveness through modeling-based training. *Organizational Dynamics, 9*, 60–77.

Roth, J. (2011). *The relationship between emotional intelligence and pastor leadership in turnaround churches* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3487845).

Roundy, C. A. (2007). *Workplace bullying: Investigating the potential link with emotional intelligence* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3302625).

Saklofske, D. H., Austin, E. J., & Minski, P. S. (2003). Factor structure and validity of a trait emotional intelligence measure. *Personality and Individual Differences, 34*, 707–721.

Salovey, P., & Birnbaum, D. (1989). The influence of mood on health-relevant cognitions. *Journal of Personality and Social Psychology, 57*, 539–551.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9*, 185–211.

Salovey, P., & Rodin, J. (1985). Cognitions about the self: Connecting feeling states to social behavior. *Review of Personality and Social Psychology, 6*, 143–166.

Salovey, P., Mayer, J. D., & Rosenhan, D. L. (1991). Mood and helping: Mood as a motivator of helping and helping as a regulator of mood. *Review of Personality and Social Psychology, 12*, 215–237.

Salovey, P., Hsee, C., & Mayer, J. D. (1993). Emotional intelligence and the self-regulation of affect. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of mental control* (pp. 258–277). Englewood Cliffs: Prentice-Hall.

Salovey, P., Mayer, J. D., Goldman, S., Turvey, C., & Palfai, T. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait meta-mood scale. In J. Pennebaker (Ed.), *Emotion,* *disclosure, and health* (pp. 125–154). Washington, DC: American Psychological Association.

Salovey, P., Stroud, L. R., Woolery, A., & Epel, E. S. (2002). Perceived emotional intelligence, stress reactivity, and symptom reports: Further explorations using the trait meta-mood scale. *Psychology and Health, 17*, 611–627.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16*, 235–239.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167–177.

Schutte, N. S., Malouff, J. M., & Bhullar, N. (2009). The assessing emotions scale. In C. Stough, D. H. Saklofske, & J. D. A. Parker (Eds.), *Assessing emotional intelligence: Theory, research, and applications* (pp. 119–134). New York: Springer.

Shank, J. C. (2012). *Emotional intelligence and educational leadership: Measuring the emotional intelligence of educational leaders and their corresponding student achievement* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3510549).

Smith, G. T., McCarthy, D. M., & Zapolski, T. C. B. (2009). On the value of homogenous construct for construct validation, theory testing and the description of psychopathology. *Psychological Assessment, 21*, 272–284.

Spencer, L. (2001). The economic value of emotional intelligence competencies and EIC-based HR programmes. In C. Cherniss & D. Goleman (Eds.), *The emotionally intelligent workplace* (pp. 45–83). San Francisco: Jossey-Bass.

Spencer, L. M. J., & Spencer, S. (1993). *Competence at work*. New York: John Wiley.

Thompson, B. L., Waltz, J., Croyle, K., & Pepper, A. C. (2007). Trait meta-mood and affect as predictors of somatic symptoms and life satisfaction. *Personality and Individual Differences, 43*, 1786–1795.

Van Rooy, D. L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior, 65*, 71–95.

Walker, A. E. (2001). *Emotional intelligence of the classroom teacher* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (9996964).

Walpole, P., Isaac, C. L., & Reynders, H. (2008). A comparison of emotional and cognitive intelligences in people with and without temporal lobe epilepsy. *Epilepsia, 49*(8), 1470–1474.

Wechsler, D. (1940). Nonintellective factors in general intelligence. *Psychological Bulletin, 37*, 444–445.

Wechsler, D. (1943). Nonintellective factors in general intelligence. *Journal of Abnormal Psychology, 38*, 100–104.

Wechsler, D. (1997). *Wechsler adult intelligence test administration and scoring manual*. San Antonio: The Psychology Corporation.

Weinberger, L. A. (2002). Emotional intelligence: Its connection to HRD theory and practice. *Human Resource Development, 1*, 215–243.

Weissman, M. (1999). *Social adjustment scale – Self-report (SAS-SR) user's manual*. North Tonawanda: Multi-Health Systems.

Woods, B. R. (2010). *The relationship between a manager's emotional intelligence and perceived leadership style* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses: Full Text (3409369).

Xue, G., Lu, Z., Levin, I. P., & Bechara, A. (2010). The impact of prior risk experiences on subsequent risky decision-making: The role of the insula. *NeuroImage, 50*(2), 709–716.

Zeidner, M., Matthews, G., & Roberts, R. (2004). Emotional intelligence in the workplace: A critical review. *Applied Psychology: An International Review, 53*, 371–399.

Zeidner, M., Roberts, R., & Matthews, G. (2008). The science of emotional intelligence. *European Psychologist, 13*(1), 64–78.

Part V

Applications of Intellectual Theory

Intelligence and Success 25

Tarmo Strenze

This chapter is about the relationship between intelligence and success. The central message of the chapter is simple: intelligent people are more successful than less intelligent people. The chapter will review the scientific evidence for that message and discuss some theoretical problems, in an attempt to show that the message is really not that simple. But at first we should probably ask: What is success?

Success can be defined as doing or achieving something that is generally considered desirable in the society. Naturally, there are many ways to be successful, and some of these ways may be in conflict with one another, so that achieving success in one field may restrict you from having success in another field. Some readers may be tempted to say that success is a purely subjective phenomenon, which each individual defines for oneself. That is certainly true, but it seems that there is usually a high degree of consensus in society as to what is desirable and what is not. Even if there are individuals who reject some form of success (for instance, claim that they do not care about money), that form of success still remains socially important and worthy of research.

T. Strenze, Ph.D. (✉)
Institute of Social Studies, University of Tartu, Estonia, Lossi 36-405, Tartu 51003, Estonia
e-mail: tarmo.strenze@ut.ee

Intelligence and Success: An Overview

So what is the evidence for the relationship between intelligence and success? Many readers would probably be convinced by specific examples of people who are known to be highly intelligent and who have achieved great success in some field. For instance, Bill Gates is rumored to have received an extremely high score on his college SAT,[1] which would mean that he also must have a very high IQ score. One can guess that Bill Gates' rise to one of the richest and most powerful men on earth must have something to do with his IQ. However, such cases never prove anything conclusively because we can always find some contrary examples. For instance, a legendary punk groupie Nancy Spungen had an IQ of 134, and still she became a drug addict and was expelled from school and, ultimately, from her own home by her own parents (Spungen 1983).

Instead of well-selected examples, we should look at the statistical relationships between intelligence and various forms of success. Such relationships have been examined in numerous studies, and it is not possible to review all of them here. This review will concentrate on meta-analyses because results from meta-analyses are

[1] The Biography Channel website. Retrieved Sep 01, 2013, from http://www.biography.com/people/bill-gates-9307520

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_25,

Table 25.1 Relationship between intelligence and measures of success (Results from meta-analyses)

Measure of success	r	k	N	Source
Academic performance in primary education	.58	4	1,791	Poropat (2009)
Educational attainment	.56	59	84,828	Strenze (2007)
Job performance (supervisory rating)	.53	425	32,124	Hunter and Hunter (1984)
Occupational attainment	.43	45	72,290	Strenze (2007)
Job performance (work sample)	.38	36	16,480	Roth et al. (2005)
Skill acquisition in work training	.38	17	6,713	Colquitt et al. (2000)
Degree attainment speed in graduate school	.35	5	1,700	Kuncel et al. (2004)
Group leadership success (group productivity)	.33	14		Judge et al. (2004)
Promotions at work	.28	9	21,290	Schmitt et al. (1984)
Interview success (interviewer rating of applicant)	.27	40	11,317	Berry et al. (2007)
Reading performance among problem children	.26	8	944	Nelson et al. (2003)
Becoming a leader in group	.25	65		Judge et al. (2004)
Academic performance in secondary education	.24	17	12,606	Poropat (2009)
Academic performance in tertiary education	.23	26	17,588	Poropat (2009)
Income	.20	31	58,758	Strenze (2007)
Having anorexia nervosa	.20	16	484	Lopez et al. (2010)
Research productivity in graduate school	.19	4	314	Kuncel et al. (2004)
Participation in group activities	.18	36		Mann (1959)
Group leadership success (group member rating)	.17	64		Judge et al. (2004)
Creativity	.17	447		Kim (2005)
Popularity among group members	.10	38		Mann (1959)
Happiness	.05	19	2,546	DeNeve & Cooper (1998)
Procrastination (needless delay of action)	.03	14	2,151	Steel (2007)
Changing jobs	.01	7	6,062	Griffeth et al. (2000)
Physical attractiveness	−.04	31	3,497	Feingold (1992)
Recidivism (repeated criminal behavior)	−.07	32	21,369	Gendreau et al. (1996)
Number of children	−.11	3		Lynn (1996)
Traffic accident involvement	−.12	10	1,020	Arthur et al. (1991)
Conformity to persuasion	−.12	7		Rhodes and Wood (1992)
Communication anxiety	−.13	8	2,548	Bourhis and Allen (1992)
Having schizophrenia	−.26	18		Woodberry et al. (2008)

r correlation between intelligence and the measure of success, *k* number of studies included in the meta-analysis, *N* number of individuals included in the meta-analysis

more reliable than results from single studies. Table 25.1 presents a list of meta-analytic correlations between IQ scores and a number of outcomes that can reasonably be designated as "success" (or lack of success). Of course, several important forms of success have never been subjected to meta-analysis and are, consequently, absent from Table 25.1; on the other hand, some forms of success have been meta-analyzed more than once, in which case I used the largest meta-analysis.

Overall, it is evident from Table 25.1 that intelligence tends to be positively correlated with desirable outcomes and negatively correlated with undesirable outcomes. Let us take a closer look at some of the outcomes in Table 25.1. One of the most classical outcomes of intelligence is academic performance, typically measured by grade point average or a specific academic test. It is classical because IQ testing was originally invented to predict academic success of students, so it is no surprise that its correlation with intelligence is positive. However, the correlation is perhaps not as strong as one might have expected. Only among elementary school students is the correlation really noteworthy (.58); among high school and college students, it is much lower (0.24 and 0.23). This result goes against the claim of some critics (e.g., McClelland 1973) that IQ test is nothing else but a test of school learning. But why is the correlation weaker on higher educational levels? The answer probably has to do with decreasing variance: as people move from elementary education to secondary and tertiary education, less intelligent students are excluded with each transition, reducing the variance of intelligence and thereby also its correlation with academic performance.

A highly desirable form of success in the modern world is career success (or socioeconomic success). In Table 25.1, it is represented by education, occupation, income, and promotions. All these things are positively correlated with intelligence – correlation with educational attainment is among the strongest correlations in Table 25.1 (0.56), correlation with occupational attainment is also strong (0.43), and income and promotions have somewhat weaker correlations (0.20 and 0.28). These results mean that intelligent people generally occupy higher positions in society. A society with such IQ-based stratification is called meritocracy (Young 1958) and is often considered to be a fair and efficient form of society, because people are allowed to achieve positions that correspond to their talents, as opposed to being allocated to positions according to their social origin (position of parents), race, or gender. There has been quite a lot of dispute on how meritocratic contemporary western society really is. In 1994, Herrnstein and Murray published a book called *The Bell Curve* that became notorious for claiming that, in the United States, intelligence has a considerably stronger effect on various forms of success than social origin and that American society is moving toward IQ-based class system. Saunders (1997) found that the same might be true for Great Britain. Such results imply that society is rather meritocratic. However, critics have argued that these studies overestimated the effect of intelligence and underestimated the effect of social origin (Breen and Goldthorpe 1999; Fischer et al. 1996).

Another important form of success is job performance, a measure of how well a worker performs his or her work tasks. That is obviously of great relevance to organizations, and much research has been devoted to finding good predictors of job performance. Positive correlations with supervisory ratings of job performance (0.53) and work sample tests (0.38) in Table 25.1 demonstrate that intelligent people are good workers, and IQ tests are, therefore, good personnel selection devices. Indeed, some researchers believe that IQ tests are the best personnel selection devices available (Schmidt and Hunter 1998). An interesting finding is that IQ tests are better predictors of performance among cognitively complex jobs, compared to less complex jobs (Ones et al. 2005). This means that IQ tests are very useful in selecting good engineers, architects, or dentists (cognitively complex jobs according to Roos and Treiman 1980); IQ tests are less useful for selecting good dishwashers, weavers, or garbage collectors, although, even among dishwashers, it is obvious that an intelligent worker is better than a less intelligent one.

Some correlations in Table 25.1 are not quite as expected. For instance, the correlation with happiness is only 0.05. One might wonder, if intelligent people are so successful in achieving desirable goals, then how come they are not significantly happier than less intelligent people? The answer to this is simple. According to some prominent theories of happiness (see Diener et al. 1999), the personal level of subjective well-being is actually rather stable and not much dependent on life events. Happiness is like a personality trait; you either have it or not, and things you achieve in life (or fail to achieve) will not affect it very strongly.

Another surprising result in Table 25.1 is the positive correlation with anorexia nervosa (0.20). Most studies have found that intelligent people are healthier and live longer than less intelligent people (Calvin et al. 2010), so why are they more likely to contract a serious disorder like anorexia nervosa? To offer a speculative answer, we can use the evolutionary theory of intelligence developed by Kanazawa (2004). According to this theory, general intelligence is a brain function that has evolved in human evolution to deal with evolutionarily novel tasks. Take, for instance, activities like finding food, having children, and collaborating with other humans – these are all tasks that our ancestors have been solving for millions of years, and for these tasks, it is likely that specific hereditary brain mechanisms have developed that promote the successful performance of that task. But activities like getting good grades at school, making a lot of money, or being thin have just recently been invented by our society, and they do not (yet) have their own brain mechanisms. For these novel tasks, people use intelligence, which is a generic ability to solve any type of (unexpected) problems. Kanazawa notes that intelligence correlates positively with evolutionarily novel activities, but the correlation with ancient activities is zero or even negative. That is also evident in Table 25.1, which mostly lists novel school- or job-related forms of success that have the expected positive correlation with intelligence, but one of the most ancient forms of success, number of children, has a negative correlation (−0.11). As for anorexia nervosa, the desire for thinness (the basis for anorexia) is clearly a novel goal that probably takes some intelligence to achieve it.

Overall, the results in Table 25.1 present a kind of a portrait of an intelligent person with positive correlations depicting the characteristics that an intelligent person is likely to have and negative correlations depicting the characteristics he or she is not likely to have. To make sense of these correlations, it is good to have theories like the one by Kanazawa (described above) that do not concentrate on just one specific form of success, but strive to explain the whole pattern of correlations.

Genes, Intelligence, and Success

It is difficult to discuss intelligence without going into discussion about genes. Intelligence certainly would not be such a controversial subject if there was no reason to believe that IQ differences between people are, to a considerable degree, caused by genetic differences. Some researchers have suggested that the heritability of intelligence could be as high as 0.80 (Jensen 1969), but most have come up with lower estimates, somewhere around 0.50 (Devlin et al. 1997), which is still quite high. Given the substantial genetic basis of intelligence and the robust relationship between intelligence and social success, one can conclude that the difference between successful and less successful people is also genetic, to some degree. This was indeed the conclusion made by Herrnstein and Murray (1994). They argued that contemporary western society allows people to fulfill their genetic potential, which means that people can achieve as much success as their genetic IQ enables them to achieve; the social position of each individual is, thus, ultimately determined by genes, and western society evolves toward genetic hierarchy where people with “good genes” live in luxury and people with “bad genes” struggle to survive. That system is further solidified by assortative mating, the tendency for people to marry and have children with partners of similar IQ. Children of intelligent and rich parents have, thus, a double advantage – they inherit their parents’ IQ as well as their resources – children of less intelligent poor parents, however, are handicapped on both accounts (Herrnstein and Murray 1994).

Is there any reason to believe that contemporary society could be such a genetic caste society? Behavioral genetic research has found that almost all human characteristics and behaviors have some genetic basis. In addition to intelligence, it has been found that criminality, alcoholism, and smoking also have a genetic component (Malouff et al. 2008). The same is true about the main social status indicators – education, occupation, and income. The heritability coefficient of education is about 0.50 (Rowe et al. 1999; Silventoinen et al. 2000), heritability of occupation is about 0.40 (Tambs et al. 1989; Plomin and Bergeman 1991), and heritability of income is about 0.30 (Taubman 1976; Rowe et al. 1999). This means that the similarity of parents and children in terms of social status is partly due to the genetic transmission of characteristics that foster (or hinder) status attainment. These numbers are interesting, but it should be noted that the heritability values of status characteristics are still much lower than the heritability values of physical characteristics. The heritability of height, for instance, is about 0.75 (Silventoinen et al. 2000; Benyamin et al. 2005). So, the diagnosis of "genetic caste society" must be an overstatement. There is probably some correspondence between the genetic structure of human population and the social structure of human society, but not any genetic IQ castes with impenetrable borders.

And what is the role of intelligence in the genetic transmission of social status? Given that intelligence is heritable, Herrnstein and Murray made an automatic conclusion that intelligence must be the characteristic that plays the central role in the intergenerational transmission of genetic advantages. But that conclusion may have been premature. Bowles and Gintis (2002) have calculated that the role of "IQ genes" in the parent–child similarity of social status is actually quite small. They do not deny that social status is heritable, but they claim that it is so mostly due to other genetic characteristics, like race, health, or personality. This conclusion by Bowles and Gintis rests on sophisticated calculations, which I have never seen anybody else make. I would treat their conclusion with some caution until their method has found more acceptance.

History, Intelligence, and Success

The evidence discussed so far has come exclusively from contemporary western societies. But what about earlier historical periods and less developed societies? Do these societies also have intelligent people on top? There is, of course, no direct evidence on the intelligence of people from earlier than the twentieth century. But the general opinion seems to be that earlier historical periods mostly did not allow intelligent people to get ahead in society. These societies had a rigid class system, and a person born to lower ranks had no opportunity to raise to upper ranks, no matter how intelligent he or she was. According to *The Bell Curve*, western societies really started to become more meritocratic only in the middle of the twentieth century. Around that time, the educational system became more democratic, and colleges were opened up to intelligent youth, irrespective of their social background. At the same time, the occupational system became more complex with a lot of new cognitively demanding jobs requiring intelligent workers. These two historical developments – increasing openness and complexity – are the main social factors that created the positive correlation between intelligence and career success, according to Herrnstein and Murray (1994).

This scenario sounds convincing, but it has been criticized on several grounds. Some authors have presented evidence showing that intelligent people were, in fact, able to be successful in earlier historical periods. Weiss describes the towns of the sixteenth- to seventeenth-century Germany where a lot of young men from modest social background were able to work themselves into higher positions (Weiss 1995). Weiss speculates that these men probably had higher than average IQ, given that they had no other advantage that would explain their rise. Botton describes how Napoleon changed the nineteenth-century French army so that new officers were recruited and promoted on the basis of their talent, rather than social background (Botton 2004). Again, we can speculate that this new system increased the correlation between intelligence and rank in the army (and perhaps in society, more generally). Adkins and Guo go much further back in time to claim that the positive effect of genetic characteristics (such as intelligence) on success was probably the strongest in the archaic society of hunter-gatherers, before the emergence of private property and desire to pass it on to children (Adkins and Guo 2008). Of course, "success" had a completely different meaning back then, but it depended entirely on individual ability.

Another line of criticism against *The Bell Curve* concerns the claim of increasing meritocracy during the twentieth century. A number of studies have tried to test this claim, and most have failed to find the strengthening of the IQ–success correlation, predicted by *The Bell Curve* (Hauser and Huang 1997; Bowles et al. 2001; Strenze 2007). All these studies have used data collected over several decades (mostly starting with the 1960s), and they have not found any signs of the IQ–success relationship getting

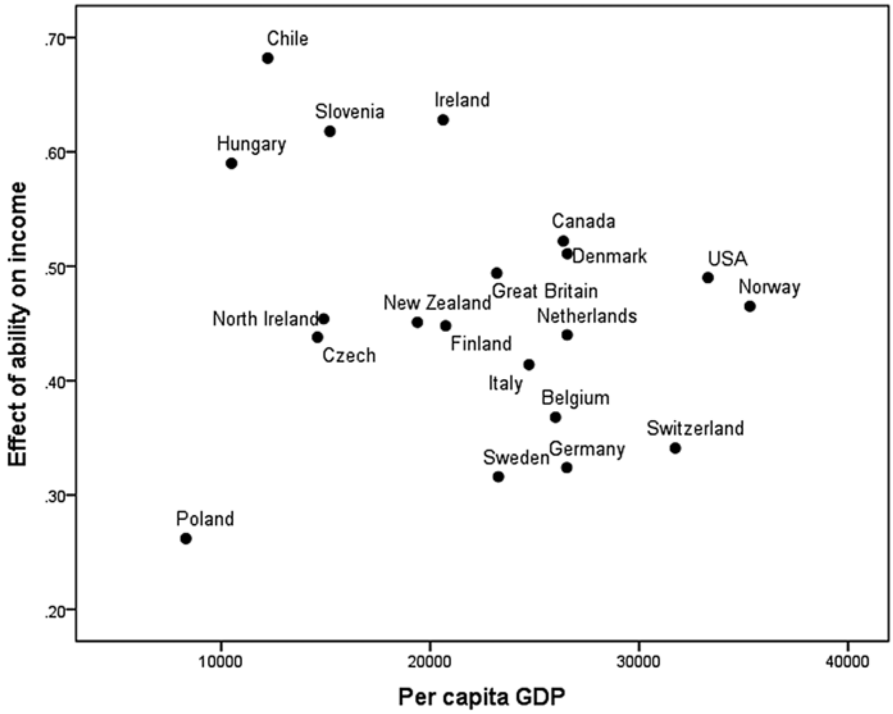


Fig. 25.1 The relationship between IQ–success relationship in a country (vertical axis) and economic development of a country (horizontal axis)

stronger during that time. Does it mean that the thesis of increasing meritocracy is wrong? Perhaps, but it must be said that these studies have almost exclusively relied on data from the second half of the twentieth century, beginning with the 1960s; there is little usable data from the first half of the century. According to *The Bell Curve*, however, the most radical change took place somewhere in the middle of the century (at least in the Unites States) – so it is possible that the available data are simply too late for the change we are looking for. But at least it seems relatively safe to say that after 1960–1970 there has been little change in the correlation between IQ and success.

An alternative way to address the same issue is to compare data from different countries to see if more developed countries have a stronger relationship between intelligence and success – that would support the idea that societies become more meritocratic as they evolve from traditional into industrial and postindustrial. A sophisticated cross-national analysis of that kind has not yet been conducted because of lack of data. But as a preliminary gauge, take a look at Fig. 25.1 that presents a simple scatterplot based on data from the International Adult Literacy Survey (IALS). IALS is a cross-national survey, conducted in 1994–1998, that measured the literacy ability of adults in 20 countries; it also included data on the career success of these adults. I calculated for each country the effect (regression coefficient) of literacy ability on income (see Strenze 2013, for details) – that effect is presented on the vertical axis of Fig. 25.1. The horizontal axis of Fig. 25.1 is the 1995 per capita gross national product (GDP), a measure of societal development. Based on the reasoning offered above, one would expect to find a positive relationship between GDP and ability–income correlation, but in fact the relationship in Fig. 25.1 is negative. The more developed countries with higher GDP, like Norway or Switzerland, tend to exhibit lower correlations between people's ability and income than the less developed countries like Chile or Hungary. Of course, the data used in Fig. 25.1 is far from perfect, and the number of countries is too small to draw any ironclad conclusions. However, a similar result was obtained by Psacharopoulos and Patrinos (2004) as they compared the relationship between education and income in nearly 100 countries and found that the relationship is stronger in less developed countries. That supports the impression that, among the societies that exist today, less developed societies are the ones where people with higher ability (and education) get better financial rewards.

Based on these results, we can piece together a speculative scenario of the history of the IQ–success relationship. The primitive society of our ancestors was presumably rather meritocratic as each person had to earn one's place in the tribe using one's own abilities and nobody got any help from their "rich daddy." As human society grew more complex, large inequalities between social groups emerged (think of slaves and citizens in ancient Rome or peasants and aristocrats in medieval Europe), and most people were destined to live in the social class of their parents – intelligence probably had little effect on people's life in these societies. These rigid class boundaries started to break down with the advent of industrial society, as democratic values became prevalent and there was increased demand for able workers – that created an opportunity for intelligent people to move up in the social ladder. This process apparently reached its apex in the middle of the twentieth century (in western societies) when the final push toward liberalization of educational and occupational market took place. But after that, it seems, the relationship between intelligence and career success has stayed the same or even declined, possibly due to the tendencies in postindustrial welfare society to reduce inequality and competition.

Conclusion

This short chapter was about the social consequences of intelligence. We saw that high IQ generally helps people to achieve numerous desirable outcomes and to keep away from undesirable ones (but there are also some interesting exceptions to that). The relationship between intelligence and success is partly based on genetics, as intelligence is itself partly genetic characteristic. But on the other hand, the relationship is partly based on societal context, as only certain social conditions allow intelligent people to fulfill their potential.

As we think about intelligence and success, we must remember that the scientific question about the relationship between intelligence and success is closely connected to other scientific questions about intelligence and, most importantly, to the following question: What is it that IQ tests really measure? This chapter was based on the implicit assumption that IQ tests are reasonably good measures of general cognitive ability, but not all social scientists would agree with that. To make sense of the correlations between intelligence and success, one must have a view on IQ testing and on the nature of intelligence, in general, which is why I now direct the reader to other chapters of this book where these related topics are discussed.

References

Adkins, D. E., & Guo, G. (2008). Societal development and the shifting influence of the genome on status attainment. *Research in Social Stratification and Mobility, 26*, 235–255.

Arthur, W., Barret, G. V., & Alexander, R. A. (1991). Prediction of vehicular accident involvement: A meta-analysis. *Human Performance, 4*, 89–105.

Benyamin, B., Wilson, V., Whalley, L. J., Visscher, P. M., & Deary, I. J. (2005). Large, consistent estimates of the heritability of cognitive ability in two entire populations of 11-year-old twins from Scottish mental surveys of 1932 and 1947. *Behavior Genetics, 35*, 525–534.

Berry, C. M., Sackett, P. R., & Landers, R. N. (2007). Revisiting interview-cognitive ability relationships: Attending to specific range restriction mechanisms in meta-analysis. *Personnel Psychology, 60*, 837–874.

Botton, A. (2004). *Status anxiety*. Hamish Hamilton: Penguin.

Bourhis, J., & Allen, M. (1992). Meta-analysis of the relationship between communication apprehension and cognitive performance. *Communication Education, 41*, 68–76.

Bowles, S., & Gintis, H. (2002). The inheritance of inequality. *Journal of Economic Perspectives, 16*, 3–30.

Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature, 39*, 1137–1176.

Breen, R., & Goldthorpe, J. H. (1999). Class inequality and meritocracy: A critique of Saunders and an alternative analysis. *British Journal of Sociology, 50*, 1–27.

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B. A., Der, G., Leckenby, N., & Batty, G. D. (2010). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology*, 1–19.

Colquitt, J. A., LePine, J. A., & Noe, R. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*, 678–707.

Devlin, B., Daniels, M., & Roeder, K. (1997). The heritability of IQ. *Nature, 388*, 468–471.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well being: Three decades of progress. *Psychological Bulletin, 125*, 276–302.

DeNeve, K. M. & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. Psychological Bulletin, 124, 197–229.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*, 304–341.

Fischer, C. S., Hout, M., Jankowski, M. S., Lucas, S. R., Swidler, A., & Voss, K. (1996). *Inequality by design. Cracking the Bell Curve myth*. Princeton, NJ: Princeton University Press.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*, 575–607.

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management, 26*, 463–488.

Hauser, R. M., & Huang, M.-H. (1997). Verbal ability and socioeconomic success: A trend analysis. *Social Science Research, 26*, 331–376.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve. Intelligence and class structure in American life*. New York: Free Press.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72–98.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1–123.

Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology, 89*, 542–552.

Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review, 111*, 512–523.

Kim, K. H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education, 16*, 57–66.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86*, 148–161.

Lopez, C., Stahl, D., & Tchanturia, K. (2010). Estimated intelligence quotient in anorexia nervosa: A systematic review and meta-analysis of the literature. *Annals of General Psychiatry, 9*, 40.

Lynn, R. (1996). *Dysgenics. Genetic deterioration in modern populations*. London: Praeger.

Malouff, J. M., Rooke, S., & Schutte, N. S. (2008). The heritability of human behavior: Results of aggregating meta-analyses. *Current Psychology, 27*, 153–161.

Mann, R. D. (1959). A review of the relationships between personality and performance in small groups. *Psychological Bulletin, 56*, 241–270.

McClelland, D. C. (1973). Testing for competence rather than for intelligence. *American Psychologist, 28*, 1–14.

Nelson, J. R., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research & Practice, 18*, 255–267.

Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Cognitive ability in personnel selection decisions. In A. Evers, N. Anderson, & O. Voskuijl (Eds.), *The Blackwell handbook of personnel selection* (pp. 143–173). Malden, MA: Blackwell.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on environmental measures. *Behavioral and Brain Sciences, 14*, 373–427.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338.

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics, 12*, 111–134.

Rhodes, N., & Wood, W. (1992). Self-esteem and intelligence affect influenceability: The mediating role of message reception. *Psychological Bulletin, 111*, 156–171.

Roos, P. A., & Treiman, D. J. (1980). DOT scales for the 1970 census classification. In A. R. Miller, D. J. Treiman, P. S. Cain, & P. A. Roos (Eds.), *Work, jobs, and occupations: A critical review of the "Dictionary of Occupational Titles"* (pp. 336–389). Washington, D.C.: National Academy Press.

Roth, P. L., Bobko, P., & Farland, L. A. (2005). A Meta-analysis of work sample test validity: Updating and integrating some classic literature. *Personnel Psychology, 58*, 1009–1037.

Rowe, D. C., Vesterdal, W. J., & Rodgers, J. L. (1999). Herrnstein's syllogism: Genetic and shared environmental influences on IQ, education, and income. *Intelligence, 26*, 405–423.

Saunders, P. (1997). Social mobility in Britain: An empirical evaluation of two competing explanations. *Sociology, 31*, 261–288.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical findings of 85 years of research findings. *Psychological Bulletin, 124*, 262–274.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37*, 407–422.

Silventoinen, K., Kaprio, J., & Lahelma, E. (2000). Genetic and environmental contributions to the association between body height and educational attainment: A study of adult Finnish twins. *Behavior Genetics, 30*, 477–485.

Spungen, D. (1983). *And I don't want to live this life*. New York: Fawcett Crest.

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin, 133*, 65–94.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*, 401–426.

Strenze, T. (2013). Allocation of talent in society and its effect on economic development. *Intelligence, 41*, 193–202.

Tambs, K., Sundet, J. M., Magnus, P., & Berg, K. (1989). Genetic and environmental contributions to the covariance between occupational status, educational attainment, and IQ: A study of twins. *Behavior Genetics, 19*, 209–222.

Taubman, P. (1976). Earnings, education, genetics, and environment. *Journal of Human Resources, 11*, 447–461.

Weiss, V. (1995). The emergence of a cognitive elite. Comment on The Bell Curve. *Mankind Quarterly, 35*, 373–390.

Woodberry, K. A., Giuliano, A. J., & Seidman, L. J. (2008). Premorbid IQ in schizophrenia: A meta-analytic review. *The American Journal of Psychiatry, 165*, 579–587.

Young, M. (1958). *The rise of meritocracy 1870–2033*. London: Penguin.

The Use of Intelligence Tests in the Diagnosis of Specific Reading Disability

26

Nancy Mather and Deborah Schneider

The Use of Intelligence Tests in the Diagnosis of Specific Reading Disability

Specific reading disability has been the subject of formal academic inquiry for over a century. Throughout this period, intelligence tests have played a central but constantly evolving role in the evaluation and diagnosis of this disorder. Within this chapter, we do not examine specific intelligence tests per se but rather their historic and contemporary uses in the identification and diagnosis of specific reading disability. We discuss (a) the current definition of reading disability, (b) a brief historical perspective on the use of intelligence tests to identify and diagnose specific reading disability, (c) present-day methods of diagnosing specific reading disability, (d) specific cognitive constructs and their relevance to the accurate diagnosis of reading disability, and (e) the future use of intelligence tests in the identification and diagnosis of a specific reading disability, often referred to as dyslexia. Although we view these terms as being synonymous, throughout this chapter, we will most often use the term specific reading disability, as some object to the term "dyslexia," viewing it as a medical diagnosis, or even doubting its very existence. As Siegel and Mazabel (2013) lamented, "We do not understand why the term "dyslexia" is often viewed as if it were a four-letter word, not to be uttered in polite company" (p. 187).

N. Mather, Ph.D. (✉)
University of Arizona,
7161 East Kiva Way, Tucson, AZ 85750, USA
e-mail: nmather@u.arizona.edu

D. Schneider
Doctoral Candidate and Department Fellow
in Special Education, University of Arizona,
10814 E. Limberlost Rd, Tucson, AZ 85749, USA
e-mail: Deborah@email.arizona.edu

Definitions

Although a universally accepted definition of specific reading disability (SRD) has not been established, most contemporary definitions emphasize the following characteristics: (a) it is a neurobiological disorder that affects the mastery of sound-symbol correspondences and the fluency and automaticity of reading, writing, and spelling; (b) it is often accompanied by specific weaknesses in cognitive factors that predict poor reading and spelling performance, such as phonological awareness or processing speed; (c) it is a lifelong condition, but effective interventions can reduce its impact; (d) it can occur in individuals of any level of intelligence; and (e) other abilities are often intact and can even be advanced, thus making this type of reading disability a specific disorder. In the next sections, we discuss some of the most commonly used definitions of specific reading disability in educational and clinical practice.

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_26,

Diagnostic and Statistical Manual of Mental Disorders

The fourth edition of Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) of the American Psychological Association (APA) (1994) defined specific reading disorder (elsewhere characterized as dyslexia, specific reading disability, or reading-related learning disorder) as reading achievement "…that falls substantially below that expected given the individual's chronological age, measured intelligence, and age-appropriate education…[and] significantly interferes with academic achievement or with activities of daily living that require reading skills" (p. 48). The fifth edition of the Manual (DSM-5) (American Psychological Association 2013), however, has abandoned reading disorder as an independent diagnosis altogether, subsuming it under the diagnostic umbrella of specific learning disorder. Under this new definition, reading-related learning disorders may be present when reading and/or spelling abilities "…are substantially and quantifiably below those expected for the individual's chronological age and cause significant interference with academic or occupational performance, or with activities of daily living…" (American Psychological Association 2013, p. 67).

Mirroring the language of the Individuals with Disabilities Education Improvement Act (IDEA 2004), the APA has removed the requirement of a discrepancy between ability and achievement for the diagnosis of specific learning disorders, while notably citing "Intellectual Disability, Global Developmental Delay, uncorrected visual or auditory acuity, other mental or neurological disorders, psychosocial adversity, lack of proficiency in the language of academic instruction, or inadequate educational instruction" as exclusionary conditions (US Department of Education 2004). Though this definition eliminates the necessity of a significant discrepancy between intelligence and achievement for the diagnosis of specific reading disability, it does not preclude its use or usefulness as a diagnostic tool (Decker et al. 2013; Swanson 2013).

International Dyslexia Association

The most widely accepted definition of specific reading disability in the United States is that of the International Dyslexia Association (IDA) (Lyon et al. 2003). Despite changes in federal education law, the IDA continues to employ a definition of specific reading disability (termed *dyslexia*) in which an unexpected discrepancy between reading achievement and cognitive ability is a characteristic feature: "[Dyslexia] is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction" (International Dyslexia Association 2007). This definition, though employed by the National Institute of Child Health and Human Development (NICHD), is controversial. The major reason that it has been criticized is for purportedly undue emphasis on the phonological component of dyslexia. A number of authors have argued that the phonological deficit model is inadequate for explaining all cases of reading disorder (Peterson and Pennington 2012; Snowling and Hulme 2012) and that its importance has been overstated (Swanson et al. 2003). Consequently, other advocacy and professional organizations include in their definitions additional cognitive factors that are correlates of dyslexia.

British Dyslexia Association

The British Dyslexia Association (BDA), for example, describes dyslexia as "characterised by difficulties with phonological processing, rapid naming, working memory, processing speed, and the automatic development of skills that may not match up to an individual's other cognitive abilities" (British Dyslexia Association 2009). Although this definition of dyslexia moves beyond the phonological processing deficits identified by the IDA, it maintains two central aspects of dyslexia: (a) that the disorder is caused by a deficiency in certain cognitive processes and (b) that the reading performance is not commensurate with the person's other abilities.

Diagnosis of Specific Reading Disability

To be effective, remedial instruction in reading must be preceded by careful diagnosis (Monroe and Backus 1937, p. 12).

Historically, intelligence tests have been used to diagnose specific reading disability by two primary means: (a) the identification of significant discrepancies between an individual's assessed ability and academic performance (an ability-achievement discrepancy) and (b) the detection of significant performance variations among cognitive and academic abilities within an individual (often referred to as intraindividual variations or a pattern of strengths and weaknesses). In this next section, we discuss the historical antecedents and origins of these two approaches.

Early Case Descriptions

The construct of reading disability long predates recent debate regarding its definition. Children and adults with difficulties consistent with reading disability were first described in the research literature in the late nineteenth and early twentieth centuries. As an example, Kussmaul, a German neurologist, described an otherwise typically abled adult patient with what he characterized as word blindness and noted that word blindness "may exist although the power of sight, the intellect, and the powers of speech are intact" (Kussmaul 1877, p. 595). By emphasizing the specificity of the reading disability and the intact state of other areas of cognitive functioning, Kussmaul gave birth to the construct of specific reading disability (Hallahan and Mercer 2002).

In 1896, Morgan used Kussmaul's term (word blindness) to describe the condition of an otherwise typically developing 14-year-old boy who would have been "the smartest lad in the school if the instruction were entirely oral" (p. 94). In this case study, Morgan remarked that the word blindness experienced by his patient "follow[ed] upon no injury or illness, but [was] evidently congenital, and due most probably to defective development of that region of the brain" (Morgan 1896, p. 1378). This observation gave rise to the idea of a congenital, biological basis for specific reading disability.

Hinshelwood (1917), an ophthalmologist and surgeon at the Glasgow Eye Infirmary, provided additional case study evidence to support the view of specific reading disability as a congenital disorder of neurobiological origin. He described a family in which six children, across two generations, suffered from varying degrees of congenital word blindness, characterized by an "inability to recognize by sight, words and letters…" (Hinshelwood 1917, p. 72). Like Kussmaul (1877) and Morgan (1896), Hinshelwood (1917) noted that his subjects suffered from no other known functional deficits and were otherwise typical in their development.

Origin of the Ability-Achievement Discrepancy Procedure

The practice of using information derived from intelligence tests to help identify specific reading disabilities is also well established in the historical literature. In 1925, Orton first used the Stanford-Binet test of intelligence to identify several students whose reading abilities were well below what would have been typical of children of comparable intelligence (Hallahan and Mercer 2002). Less than a decade later, Monroe (1932) used a variety of achievement and intelligence tests, including the Stanford-Binet, to identify students who had discrepancies between their expected abilities (per the Stanford-Binet) and their performance on measures of reading achievement.

Monroe, a psychologist and author of many of the *Dick and Jane* readers, began her academic career as a research associate for Orton, codeveloper of the Orton-Gillingham reading method. In her classic book, *Children Who Cannot Read*, Monroe (1932) recommended the use of a reading index that was designed to predict reading performance based on age and ability. The intent of this index was to quantify the concept of unexpected underachievement or reading achievement that was significantly lower than would be predicted by chronological age, mental age, and academic accomplishments in other areas.

To that end, Monroe developed a reading expectancy formula that was designed to determine if an individual's reading achievement was at a level consistent with her or his other achievements. The factors used to determine reading expectancy comprised (a) chronological age, (b) mental age (based upon the Stanford-Binet test of intelligence), and (c) current levels of performance in arithmetic computation. This expectancy score was then compared to the average level of reading achievement across several standardized assessments. Describing her expectancy formula, Monroe explained: "It seems that we are measuring a discrepancy between reading and other accomplishments, which may occur in either direction at any intellectual level" (p. 17). Upon the identification of such a discrepancy, Monroe recommended that remedial reading instruction be provided until reading performance became consistent with intellectual ability and academic accomplishments in other areas (p. 177).

Despite her endorsement of this index, Monroe noted that her method was not without problems. The most significant of these was that the original Stanford-Binet, used to calculate intelligence for the reading expectancy formula, required reading on various subtests that were not designed primarily to measure reading ability. Many students referred to Monroe were unable to read at all at the time of testing, resulting in unreliable scores and the likely under-prediction of intelligence and reading potential (Monroe 1932). Monroe and her colleague Backus (1937) observed, "Sometimes children of good general intelligence show retardation in some of the specific skills which compose an intelligence test" (p. 22). This observation was consistent with observations of Monroe's mentor, Dr. Samuel Orton (1925), who noted that "…psychometric tests as ordinarily employed give an entirely erroneous and unfair estimate of the intellectual capacity of these children" (p. 582). This obvious drawback notwithstanding, Monroe and Backus (1937) advised that all poor readers be given an intelligence test "…to determine whether the child has the mental capacity for a higher level of reading than he is actually achieving" (p. 21). Monroe and Backus (1937) added the qualification, however, that the selected test should not be too heavily weighted on one type of ability and that tests requiring considerable reading should be avoided because they would unduly penalize poor readers.

Origin of the Intraindividual Variation Procedure

In addition to developing the concept of an ability-achievement discrepancy, Monroe and Backus (1937) pioneered the use of standardized tests of intelligence to identify patterns of strengths and weaknesses in students' intellectual profiles. They observed that "[a]n analysis of the various sub-tests in the Stanford-Binet or other intelligence tests may give an indication of the child's best and poorest intellectual abilities" (p. 23). They did not, however, go so far as to advocate the use of intraindividual variations in cognitive performance as a diagnostic tool in its own right.

Travis, widely recognized in the United States as the founding father of the discipline of speech-language pathology, was among the first practitioners to examine the unique patterns of cognitive and academic strengths and weaknesses particular to individuals with reading disabilities. In a chapter written for the *Yearbook of the National Society for Studies in Education*, Travis (1935) described children of typical intelligence who struggled to learn to read or spell but performed proficiently in other areas. These children, whom he characterized as possessing a *special* disability, showed striking variations in abilities among subjects as well as notable discrepancies between measured intelligence and expected performance in subject-specific areas. Some of these children could not read, although they could comprehend material read to them; others could read proficiently but could not comprehend what they had read. Travis explained that the clearest expression of this type of special disability was an academic profile comprising consistently low scores on tests in a given subject and average or above average scores on tests in other subjects. He explained that "…in the case of a reading disability, a child might obtain scores placing him in the ninth grade in arithmetic…and in the third grade in reading. Here we would have evidence of a striking reading disability" (p. 43).

A key advantage cited by early advocates of the intraindividual variation procedure was the educationally useful information provided by students' performance on the various subtests within a comprehensive assessment battery. Perhaps this idea was best expressed by Gallagher (1966), who noted that "[t]he information provided by this patterning of abilities is much more important than [a] single mental age score…It not only establishes the fact of developmental imbalance, but it locates the areas of specific disability" that, presumably, could be addressed through targeted remediation (Gallagher 1966, p. 30). The use of intraindividual variations in the diagnosis of specific learning disabilities, however, was not without its critics. Bijou (1942), a contemporary of Travis, first observed that cognitive tests were not designed to diagnose disability through pattern analysis and might therefore be invalid for such a purpose. This observation may have been prescient, as later research (e.g., Kavale and Forness 1984; O'Donnell 1980; Proctor and Prevatt 2003) has likewise called into question the existence of a unique pattern of cognitive strengths and weaknesses that characterizes specific reading disability. Nevertheless, all historical approaches to SLD identification have emphasized an analysis of the spared or intact abilities that "stand in stark contrast to the deficient abilities" (Kaufman 2008, p, 7).

Present-Day Methods for Diagnosing SRD

The determination of a disability should definitely constitute more than a simple "exercise in arithmetic" (Willis and Dumont 2002, p.173).

Because specific reading disability falls under the more generic category of specific learning disability, federal guidelines for SLD identification within public schools also apply to this subcategory. Currently, under the Individuals with Disabilities Education Improvement Act (IDEA 2004) (US Department of Education 2004), three methods may be used for the diagnosis of SLD: (a) ability-achievement discrepancy, (b) response to intervention, and (c) alternative research-based procedures. In this next section, we discuss contemporary uses of each of these methods, as well as some of the controversies and concerns related to each.

Ability-Achievement Discrepancy

The US Office of Education first codified the use of an ability-achievement discrepancy procedure to identify and diagnose reading disabilities in 1976 (US Office of Education 1976). Under the federal regulations of the day, "[a] specific learning disability [could] be found if a child [had] a severe discrepancy between achievement and intellectual ability in one or more of several areas: oral expression, written expression, listening comprehension or reading comprehension, basic reading skills, mathematics calculation, mathematics reasoning, or spelling" (US Office of Education 1976, p. 52405). For three decades, spanning the mid-1970s to the mid-2000s, federal law required that SLD status be determined primarily by the documentation of a significant ability-achievement discrepancy. While the diagnosis of an ability-achievement discrepancy was no longer required as of the 2004 reauthorization of IDEA (legally in force as of 2006), the procedure is still widely used in American public schools today.

An ability-achievement discrepancy is diagnosed when performance in one or more areas of achievement, per standardized testing, varies significantly from values predicted by overall intellectual ability per standardized testing. In the majority of cases in which the discrepancy model is used, diagnosis of a specific learning disability and concomitant provision of special education services require several components. These include but are not necessarily limited to (a) documentation of a statistically significant discrepancy, often 1.5 standard deviations or more, between one or more areas of academic achievement and intellectual ability; (b) determination of a link between the identified discrepancy and the poor academic performance; (c) a team-based determination that an individual's educational needs would be best met through the provision of special education services; and (d) consideration of exclusionary criteria, such as intellectual disability or English language learner status (Restori et al. 2009).

Despite its long-standing use and codification in law, the diagnostic use of the ability-achievement discrepancy procedure has proven to be one of the most contentious issues in the field of learning disabilities (Swanson 2013). The ability-achievement discrepancy procedure was designed as one means by which to identify the existence of *unexpected underachievement* (i.e., achievement far below what would be predicted based on overall intelligence); it was not initially intended to be used as the primary criterion upon which to diagnose SLD – and research has not supported its use as such (Fletcher et al. 1992; Fletcher et al. 1994).

For example, in a meta-analysis of 46 studies comparing groups of struggling readers with ability-achievement discrepancies to those without discrepancies, Stuebing et al. (2002) concluded that the use of an ability-achievement discrepancy for the identification of a reading disability had weak validity. This is likely at least partially attributable to an operational mismatch between the definition of specific learning disability, which characterizes SLD as a deficit in psychological processing, and the purpose of the ability-achievement discrepancy procedure: the identification of unexpected underachievement. Other criticisms of the use of ability-achievement discrepancy of SLD identification have focused on the following:

1. The magnitude of the discrepancy tends to increase with age, and young children often lack an ability-achievement discrepancy sufficient for diagnosis, even when well behind their peers (Fletcher et al. 1992; Kavale 2002; Speece 2002). This can result in educators waiting for students to fail before referring them to special education evaluation and services. Delays in service provision caused by such *wait-to-fail* approaches can exacerbate academic difficulties and cause students to lose ground outside of their immediate areas of disability.
2. Aspects of the learning disability itself, such as impaired processing speed or poor working memory, may disproportionately impact overall performance on standardized tests of intelligence, thus diminishing the size of the discrepancy. This effect tends to compound with time, as poor reading skills impact vocabulary development and overall knowledge, decreasing overall scores and reducing the size of the discrepancy.
3. The size of the discrepancy required for diagnosis of SRD varies widely among school districts and states (Berninger 1996), resulting in the uneven provision of special education services across the nation.
4. Simple discrepancy calculations fail to account for the variability in the origin and presentation of reading difficulties (Catts et al. 2003; Swanson et al. 2006), and the identification of an ability-achievement discrepancy does little to enhance understanding of the reasons for reading failure or provide guidance for the development of an intervention plan.
5. Poor readers of varying intelligence show similar reading, spelling, language, and memory deficits (Siegel 1989). These findings call in to question the practice of denying the diagnosis of specific learning disability to students with IQ scores well below average for their age.

Clearly, rigid adherence to formula-based ability-achievement models for the diagnosis of SRD lacks empirical support and may also result in decreased attention to significant qualitative and quantitative information that can be found within intelligence tests. Furthermore, contrary to the theoretical assumptions underlying the ability-achievement discrepancy procedure, intelligence tests and achievement tests do not fall into a neat binary of abilities and achievement skills (Lohman

2006). In fact, far more overlap exists between the skills and abilities measured on tests of intelligence and tests of achievement than many realize. Lohman explained that "Most novices believe that ability is innate and achievement acquired, whereas experts see the two as different aspects of the same thing" (p. 1). Assuming an absolute distinction exists between tests of achievement and tests of intelligence is erroneous as they often measure similar skills and abilities. For example, most intelligence tests have measures of oral vocabulary, whereas achievement tests often have similar vocabulary tests that require reading. Both are measures of lexical knowledge, and performance would tend to be similar unless an individual has poor reading skills. Anastasi (1980) alluded to this phenomenon when she lamented, "If a benevolent wizard were to give me the power to eliminate four words from the tester's vocabulary, I would choose 'intelligence,' 'aptitudes,' 'abilities,' and 'achievement.' Then if a malevolent wizard were suddenly to appear and demand that I take back one word, I would chose to retain 'abilities'" (p. 1). Anastasi referred to all abilities as "developed" skills that evolve through cultural experiences and interests, suggesting that abilities and skills are simply two sides of the same coin. While cognitive abilities are grounded in biological processes, they, like the skills in which they are manifested, are developed through education and instruction. Consequently, all cognitive and achievement tests are, to a great or lesser extent, measures of developed abilities (Lohman 2006), and the diagnostic utility of the ability-achievement discrepancy is therefore questionable.

Nevertheless, the existence of a discrepancy between intelligence tests and reading achievement can provide important information in certain situations. While it is not universally the case, approximately one half of students diagnosed with SLD do have significant ability-achievement discrepancies (Kavale and Reese 1992). Furthermore, Kavale et al. (1994), in a reanalysis of ability-achievement testing data, found that ability-achievement discrepancies can be used to differentiate students with specific reading disabilities from other low achievers with a moderate degree of fidelity. Such discrepancies may be attributable to a number of factors, including natural variability in skills (e.g., verbal abilities higher than reading abilities), strengths in specific areas that are measured by intelligence tests (e.g., analytic reasoning) and weaknesses in areas that are measured by the achievement tests (e.g., decoding), or a lack of appropriate instruction or remediation in reading. Nevertheless, the detection of a significant ability-achievement discrepancy warrants further investigation, whatever the cause.

Moreover, in some students with SRD, an ability-achievement discrepancy can be attributed to the reading disability itself. For such individuals, an ability-achievement discrepancy actually operationalizes the concept of unexpected underachievement (Ferrer et al. 2010, p. 10). This is particularly true for individuals of high intelligence whose reading lags far behind their overall abilities. These students, often referred to as twice exceptional, may read at a level consistent with their peers, yet underachieve significantly in relation to their academic potential and intelligence (Mather and Gerner 2009). Although their decoding skills are significantly weaker than would be expected based on their overall intelligence, they rarely qualify for special education services using methods other than an ability-achievement discrepancy (Crepeau-Hobson and Bianco 2011). While the overall abilities of twice exceptional students often allow them to compensate for their poor reading skills, they would nevertheless benefit from targeted interventions and special education services (Peterson and Pennington 2012). Otherwise, they are likely to remain "slow, dysfluent readers" (Mather et al. 2013, p. 714), hampering their ability to achieve on a level commensurate with their overall intelligence.

Response to Intervention

In lieu of an ability-achievement discrepancy, IDEA 2004 now permits schools to use a process that determines whether or not a student responds to evidence-based intervention as part of SLD evaluation procedures. Response to intervention

(RTI) is an educational approach that is primarily concerned with general, not special, education. Two essential components of RTI are universal screening and progress monitoring using curriculum-based measurement (CBM) instruments at specified intervals throughout the academic year. Based upon their overall response to evidence-based instruction and instructional interventions, students are assigned to one of three tiers of academic support: (a) Tier 1, in which high-quality core reading instruction and interventions are provided to all students in alignment with curricular goals; (b) Tier 2, in which small group instruction and targeted interventions are provided to students who fail to make satisfactory progress through Tier 1 instruction; and (c) Tier 3, in which one-on-one instruction and targeted interventions are provided to students who fail to respond to Tier 2 reading instruction (Jimerson et al. 2007). In many districts, Tier 3 is not synonymous with special education. This tier is designed to address the needs of students who continue to struggle to meet reading goals, even when provided with small group instruction. Contrary to the traditional model of SLD, which is exclusive of simple low achievement or other conditions that might cause poor performance in the classroom (e.g., intellectual disability or English language learner status), RTI Tier 3 is inclusive of all causes of low reading achievement.

In a very real sense, RTI poses a challenge to the basic construct of learning disability as a condition in which underachievement is unexpected given an individual's overall intelligence and other academic accomplishments (Kavale et al. 2005). Vaughn and Fuchs (2003) indicated that RTI can be seen as an effort to redefine SLD such that the students with the greatest academic need, not just those with unexpected underachievement, are identified and served. Likewise, Feifer (2008) suggested that RTI is one means by which to address concerns associated with the so-called "test and place" philosophy of SLD identification and the often arbitrary numerical cutoffs of the ability-achievement discrepancy procedure. While these goals are laudable, CBM measures alone do not provide sufficient evidence to diagnose or discriminate between the various causes of poor reading. RTI should not be viewed as a method for the identification of specific reading disability, but as a preventative method of early identification and intervention for reading difficulties of all types.

While RTI approaches have several notable benefits, concerns remain. As with the ability-achievement discrepancy procedure, a major limitation of RTI models is that they do not provide information concerning the underlying cause of poor reading achievement (Mather and Kaufman 2006). Furthermore, RTI data appear insufficient for the early, accurate identification of those with specific reading disabilities, as opposed to those with other causes of reading underachievement. Recent research has suggested that the one-stage screening procedures that are often employed in RTI models result in unacceptably high rates of false positives (Fuchs et al. 2012). However, Fuchs et al. (2012) found that use of a two-stage screening procedure that included several cognitive (e.g., rapid naming, phonological processing, oral language comprehension, and nonverbal reasoning) and untimed and timed reading measures greatly improved classification accuracy. Therefore, comprehensive cognitive and achievement assessments should be considered for those students who fail to respond to Tier 2 interventions or who are suspected of having a reading disability. The results of such assessments can provide valuable information concerning the causes of underachievement, as well as an identification of an individual's strengths and weaknesses that can be used to inform Tier 3 interventions or suggest the need for more tailored special education services.

Pattern of Strengths and Weaknesses

In addition to ability-achievement and RTI approaches, IDEA 2004 provides a third option to schools for SLD identification. IDEA regulation § 300.309 (a)(2)(ii) permits consideration of a pattern of strengths and weaknesses among various aspects of cognitive ability and/or academic achievement as part of the SLD identification process. The pattern of strengths and weaknesses (PSW) approach, also referred to as the intraindividual variations or intraindividual discrepancies procedure, typically involves the administration of tests of cognitive abilities and related achievement tests. Performance on these measures is evaluated by a qualified professional for the presence of statistically significant variations or discrepancies among or between areas of performance that might be relevant to the diagnosis of a specific reading disability. As with the ability-achievement procedure and the RTI approach, a determination must also be made that the underachievement is not primarily the result of a factor other than a learning disability.

A principal advantage of the PSW approach is the educationally useful information provided by the assessments. According to the Standards for Educational and Psychological Testing published by the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) (1999), the PSW approach allows a qualified examiner to interpret a variety of skills and abilities in the context of the broader assessment and other educational data, permitting the identification of "problem[s] that may not be apparent if scores on different kinds of tests are combined" (p. 123). For example, an examiner might note that an individual performs poorly on timed tests of reading and also has slow processing speed scores relative to her or his performance in other areas. While such a finding might be indicative of a specific reading disability, it would also be useful in selecting appropriate accommodations (e.g., provision of additional time for assignment completion) and targeted remediation (e.g., an intervention to improve reading speed).

No universally agreed upon process exists for the evaluation of cognitive strengths and weaknesses, and disagreements abound among practitioners. Nevertheless, the following practices are foundational to the accurate and effective clinical use of the PSW approach: (a) the use of data from a wide variety of sources; (b) the use of data analytic techniques specifically designed for the identification of relevant patterns; (c) the use of instruments and assessments with a high degree of predictive and diagnostic validity; and (d) the use of sound, empirically based decision-making strategies (Schultz et al. 2012). Likewise, when selecting among assessments for a PSW evaluation, evaluators should first determine (a) the areas of concern, (b) the age of the referred student, (c) the student's grade in school, and (d) which cognitive abilities and processes are most likely to be related to the student's particular educational challenges (McGrew and Wendling 2010).

Two of the most common approaches to the clinical evaluation of patterns of strengths and weaknesses are the discrepancy/consistency approach (D/C) (Naglieri 1999) and the cognitive hypothesis testing (CHT) approach (Hale and Fiorello 2004). The D/C approach (Naglieri 1999) involves the analysis of cognitive processes using the planning, attention, simultaneous, and successive (PASS) theory as operationalized by the first (Naglieri and Das 1997) and second (Naglieri et al. 2014) editions of the Cognitive Assessment System (CAS), as well as the evaluation of observational data and data from achievement testing. Following data collection, CAS, achievement, and observational data are systematically analyzed to determine whether relationships exist between weaknesses in cognitive processing and weaknesses in academic achievement. The results of this analysis are then used to inform educational planning and identify appropriate instructional interventions (Schultz et al. 2012).

The CHT approach involves the selection of cognitive tests based upon empirical evidence that the cognitive factors to be evaluated are likely to be associated with the specific academic problems presented by an individual student (Schultz et al. 2012). In this model, data concerning areas of academic difficulty are typically derived from RTI, descriptions of the presenting problem(s), and student case histories (Schultz et al. 2012). Cognitive testing data are collected using standardized instruments rooted in CHC theory (e.g., Woodcock-Johnson Tests of Cognitive Ability IV (Schrank et al., 2014)). As with the D/C approach, data resulting from CHT approach assessments are analyzed to identify the connections between academic difficulties and areas of cognitive weaknesses. Findings are then used to target educational interventions to a student's individual needs, building on strengths and remediating weaknesses. As an example, a student with low processing speed would likely benefit from extended time, but remediation of phonics skills would be unnecessary.

Despite its previously outlined advantages, the PSW approach is not without limitations. The most important among these is a lack of validated diagnostic criteria or well-established guidelines for the use of the procedure. Furthermore, a specific pattern of strengths and weaknesses that confirms or disconfirms the presence of SRD has not been identified. In general, PSW methods do not (a) address the fact that cognitive and achievement tests often share similar content, such as measures of vocabulary; (b) recognize that error is inherent in all measurement and may lead to an unacceptably high level of false positives; or (c) specify the magnitude of the difference among abilities necessary for the diagnosis of SRD (Stuebing et al. 2012). Furthermore disagreements exist regarding the use of an intraindividual versus an interindividual or normative approach. When using an intraindividual approach, variations within an individual's unique profile are considered in determining the existence of a reading disability, whether or not achievement falls below an absolute threshold (which would be relevant in a case of a twice exceptional individual who may have superior intelligence, but reading in the low average range). When using a normative approach, however, achievement must fall below a predetermined threshold for a diagnosis of SRD to be made. Regardless of which approach is taken, an evaluator must still use clinical judgment when diagnosing SRD or developing an individualized program of education or intervention.

Cognitive Factors and Processes Associated with Specific Reading Disability

One factor alone is often not sufficient to inhibit the act of reading; each case must be investigated with the hopes of locating as many contributing factors as follows (Monroe and Backus 1937).

The cognitive-psychological model is a widely used approach to understanding the nature and etiology of specific reading disability. This model attempts "to explain both the consistencies and inconsistencies within performance" (p. 700) through psychological study, often including assessments designed to identify the cognitive factors and processes associated with particular patterns of performance (Mather et al. 2013). A number of cognitive factors and processes are associated with reading disability. Understanding those factors and processes, and how weaknesses in one or more of them can contribute to reading difficulties, is useful to both the diagnosis and remediation of specific reading disability. While research does not support the assertion that there is a single cognitive profile indicative of SRD, analysis of specific cognitive factor scores derived from standardized cognitive batteries can provide evidence of a pattern of strengths and weaknesses indicative of a reading disability, while providing a direct link to intervention hypotheses (Decker et al. 2013). In this next section, we discuss two theoretical models of cognitive abilities and processes, as well as several specific cognitive factors that have been associated with reading disability.

Planning, Attention, Simultaneous, and Successive (PASS) Theory

Modern neurological science supports the conception of the human brain as an organ comprised of differentially functioning but interdependent regions (Das 2009; Das et al. 1994; Naglieri and Otero 2011). The planning, attention, simultaneous, and successive (PASS) theory of intelligence is heavily indebted both to modern neurological science and to Luria's (1966) pioneering work on the modularization of the brain and psychological processes. PASS theory hypothesizes the existence of three major functional units of the brain: attention, simultaneous and successive processing, and planning. The attention function of the brain involves maintaining optimal levels of arousal, focusing attention on salient stimuli, and resisting distractions; the simultaneous and successive processing functions of the brain involve integrative and serial information processing, respectively; and the planning function of the brain involve executive functioning and the organization and control of behavior (Das 2002; Naglieri and Otero 2011). A major factor distinguishing PASS theory from other models of human intelligence is its emphasis on defining human abilities on the basis of cognitive processes rather than models that focus on test content (e.g., verbal, nonverbal), modality (e.g., visual, auditory), or other attributes such as memory and reasoning.

The cognitive processes theorized in the PASS theory were first operationalized in the Cognitive Assessment System (CAS) (Naglieri and Das 1997) and more recently in the CAS-2 (Naglieri et al. 2014), two psychometric instruments designed to measure cognitive processes in each of the three functional areas described by Luria. The instrument provides scores for planning, attention, simultaneous, and successive cognitive processes, as well as a full scale (Naglieri and Das 1997). Research suggests that the CAS is well suited to the differential diagnosis of learning, attentional, and cognitive disabilities (Das 2002; Naglieri and Conway 2009; Naglieri et al. 2004). In fact, the CAS has shown the ability to discriminate between children with attention deficit/hyperactivity disorder (ADHD), children with a specific reading disability (SRD), and children in general education with a high degree of fidelity (Naglieri et al. 2004). In research performed by Naglieri et al. (2004), children with SRD showed deficits in successive processing relative to their overall PASS scores, whereas children with ADHD obtained lower planning and, to a lesser extent, attention scores than on the simultaneous and successive scales. The calculated effect sizes for the differences in performance among the groups ranged from moderate to large (Naglieri et al. 2004), suggesting that the CAS can serve as a valuable tool in the diagnosis of specific reading disabilities.

PASS theory has also been linked to academic interventions. The book *Helping Children Learn* (Naglieri and Pickering 2010) provides cognitive strategies that address specific PASS weaknesses related to academic problems. For example, if a student is having problems working with sequentially presented information, as evidenced by low successive and spelling scores, a strategy such as chunking can be used to teach the student how to manage academic tasks that place large demands on memory. Furthermore, researchers (Haddad et al. 2003; Iseman and Naglieri 2011; Naglieri and Gottling 1995, 1997; Naglieri and Johnson 2000) have demonstrated that the use of PASS-based planning strategies can improve academic performance. Elementary school students using the PASS Remedial Program (PREP; Das 1999) have shown improvements in reading skills (Boden and Kirby 1995; Carlson and Das 1997; Das et al. 1995, 2000; Parrila et al. 1999).

Cattell-Horn-Carroll Theory

Another prominent model of cognitive abilities is Cattell-Horn-Carroll (CHC) theory (Carroll 1993; Cattell 1941; Horn 1985; Horn and Noll 1997: McGrew and Flanagan 1998). This model has proven particularly influential in the fields of educational psychology and psychoeducational assessment (Konold et al. 2003; McGrew 2005) and has been operationalized in the Woodcock-Johnson Tests of Cognitive Abilities (WJ IV COG) (Schrank et al., 2014). In CHC theory, nine broad stratum abilities and over seventy narrow abilities are described. The broad stratum abilities include the following:

1. Crystallized intelligence (Gc): previously acquired knowledge; the ability to retrieve, apply, and communicate that knowledge; and the ability to make reasoned judgments based upon that knowledge
2. Fluid intelligence (Gf): the ability to reason, form judgments, acquire novel concepts, and solve problems by applying novel concepts or procedures
3. Quantitative reasoning (Gq): the ability to understand and acquire quantitative concepts, engage in quantitative reasoning, identify numerical and quantitative relationships, and solve numerical and quantitative problems
4. Reading and writing ability (Grw): the abilities to perform tasks requiring word reading and spelling

5. Short-term working memory (Gwm): the ability to hold and manipulate novel information in one's immediate consciousness
6. Long-term storage and retrieval (Glr): the ability to store and fluently retrieve previously learned associations
7. Visual processing (Gv): the ability to identify, store, recall, understand, and analyze visual patterns and representations
8. Auditory processing (Ga): the ability to identify, understand, and discriminate sounds, including speech sounds
9. Processing speed (Gs): the ability to perform cognitive tasks with speed, fluency, and automaticity

Several of the broad and narrow cognitive abilities described in CHC theory have particular relevance to the identification and diagnosis of specific reading disability. These include the broad domain of crystallized knowledge (Gc) and three of its narrow component abilities: language development, lexical knowledge, and listening ability (McGrew and Wendling 2010). Memory span and working memory, inclusive of induction and general sequential reasoning, are also correlated to reading performance, as are phonological awareness (subsumed under Ga) in processing speed (Gs) and rapid automatized naming (subsumed under Glr) (McGrew and Wendling 2010). Therefore, an evaluator should consider the selection of tests designed to measure these key abilities when evaluating an individual for the presence or absence of a specific reading disability.

Cognitive Abilities Related to Specific Reading Disability

Individual variations in cognitive abilities have been shown to be important mediators of academic outcomes (Swanson and Hoskyn 1998). Both children and adults with identified reading disabilities have cognitive deficits that are pervasive across the life span. These include poor phonological awareness (Wagner and Torgesen 1987), slow naming speed (Bowers and Wolf 1993; Wolf and Bowers 1999), poor verbal memory (Swanson 2013), and slow processing speed (Feldmann et al. 2004; Kail 1991; Kail and Hall 1994; Kail et al. 1999. Discrepant performance among cognitive abilities can provide important information about which areas are most likely to require remediation and which interventions are most likely to be beneficial for an individual student.

Phonological Awareness

Phonological awareness comprises the ability to perceive and identify individual speech sounds as well as the ability to perceive and understand the phonological structure of spoken words. These abilities are subsumed under auditory processing (Ga) in the CHC model. While the capacity to produce oral language is inherent, reading and writing, and the skills upon which they are dependent, must be learned. In order to read and write, an individual must develop phonological awareness and learn to distinguish among the discrete sounds of human speech, finally mapping those sounds onto the symbols with which they are encoded in written language. Furthermore, in order to read and write efficiently, this process must occur with automaticity and ease. Phonological awareness is positively correlated to reading ability (Ehri et al. 2001), and a significant majority of individuals with specific reading disabilities perform poorly on these measures (Catts and Kamhi 2005; Wagner and Torgesen 1987). Deficits in the areas of phonological awareness are major indicators of specific reading disorder, particularly in cases in which other cognitive abilities are intact or largely intact. Individuals who are diagnosed with significant weaknesses in phonological awareness should be offered interventions designed to improve skills in these areas, followed by, or in concert with, systematic instruction in phoneme-grapheme relationships.

Rapid Automatized Naming

Tests of rapid automatized naming (RAN) measure the rapid retrieval of continuously presented, highly familiar symbols of a single type (e.g., objects, colors, letters, or numbers). Tests of rapid alternating stimuli (RAS), a measure closely related to RAN, involve the rapid retrieval of a combination of these common symbol types.

Although the cognitive abilities underlying performance on RAN tests are not fully understood, such tasks appear to capture a unique cognitive ability or set of abilities. Performance on RAN tasks is highly correlated with reading ability (Powell et al. 2007; Wolf and Bowers 1999), and poor RAN performance is a cognitive marker for specific reading disability (Georgiou and Parrila 2013) that predicts reading ability independent of other cognitive abilities, such as phonemic awareness (Powell et al. 2007). Measures of RAN capture a cognitive skill or set of skills whose importance increases with age, most likely because of increased demands on visual word recognition and automaticity (Vaessen and Blomert 2013). These measures appear to be more related to facility in reading irregular words than to phonic skills (Abu-Hamour et al. 2012). Moreover, RAN tasks have been shown to reliably predict future reading ability in preliterate children (Furnes and Samuelsson 2011; Lervåg and Hulme 2009), making RAN a useful screening tool for the early identification of specific reading disability.

Working Memory

In addition to phonologically related processes, memory-related processes also appear to contribute to specific reading disability (Crews and D'Amato 2009; Jacobson et al. 2011; Swanson et al. 2009). Working memory comprises the ability to hold and manipulate newly acquired information and concepts in immediate consciousness. It is relevant to every aspect of learning and areas of academic achievement, including decoding, encoding, and reading comprehension. In a meta-analysis of the results of 43 studies conducted between 1963 and 2006, Swanson et al. (2009) found that children with reading disabilities performed poorly on tasks that required simultaneous processing of information, nonword repetition, as well as the immediate recall of letter, number, and word strings. Poor performance in such tasks is associated with deficits in working memory, particularly in the area of verbal short-term memory. Students diagnosed with significant deficits in working memory may require explicit instruction in storing and processing sounds in spoken words; storing and processing letters, letters in written words, and written words; naming orthographic symbols; and integrating the visual appearance of written words with internally stored representations (Berninger and Swanson 2013). They may also benefit from explicit instruction in memorization strategies such as chunking, repetition, rehearsal, and the use of mnemonic devices, as well as the use of memory aids, including planners, organizers, and note cards.

Processing Speed

Processing speed, or Gs in the CHC model, comprises the rapidity, fluency, and efficiency with which an individual performs cognitive tasks. This cognitive ability appears to be another important mediator and correlate of reading disability in both children and adults (Evans et al. 2002; Gregg et al. 2006; Kail 1991; Kail et al. 1999; Konold et al. 2003; Shanahan et al. 2006; Willcutt et al. 2005). Even among individuals who have strong decoding skills, poor processing speed has been shown to hamper reading efficiency and fluency (Jacobson et al. 2011). Poor performance on measures of processing speed is associated with reduced reading rate and difficulty in fluently decoding words with irregular spelling patterns (e.g., yacht). Furthermore, impaired processing speed may hamper the ability of students to develop fluent oral reading (Jacobson et al. 2011). When compounded with deficits in phonological awareness and/or working memory, deficits in processing speed can further impair reading ability. Individuals with such deficits often require interventions and strategies designed to build reading speed, as well as additional time for assignments requiring reading.

Multiple Deficit Model

Some abilities measured by comprehensive cognitive and achievement batteries are highly correlated; this is particularly the case in the various abilities that comprise the capacity to read and write with fluency and accuracy. The understanding that many learning disorders are multifactorial in origin and etiology is called the multiple deficit model (Pennington 2006). As an example, Konold et al. (2003) assessed how children performed on measures of phonological awareness, vocabulary and listening comprehension, processing speed, and short-term memory. Among children with the strongest overall reading ability, performance tended to be high across all four measures; among children with the poorest overall reading ability, however, performance tended to be low across all four measures. There is also substantial evidence that reading disability, slow processing speed, and attention-deficit/hyperactivity disorder (ADHD), a disorder in which a high degree of distractibility and impulsivity is characteristic, have a high degree of comorbidity (McGrath et al. 2011; Willcutt et al. 2001; Willcutt et al. 2010). Furthermore, a mounting body of research suggests that speech and language disorders, phonological processing deficits, and specific reading disabilities co-occur at levels far higher than can be accounted for by chance (Willcutt et al. 2010).

Nevertheless, some of the abilities measured on intelligence tests are not highly correlated. For example, in homogenous samples of young adults, Horn and Blankson (2012) reported, "… measures in which there is much emphasis on speediness correlate near zero, perhaps negatively, with tests that require solving difficult problems" (p. 91). Thus, it is not uncommon for an individual with a specific reading disability to have high scores on measures of problem solving and reasoning but low scores on speeded tests. Such variations within individual profiles should not be seen as diagnostic of reading disability in themselves, but should be considered in relation to their relevance to specific areas of academic difficulty. A sound, empirically based decision-making procedure should then be used to determine whether or not special education services would be likely to benefit the child.

The Future Use of Intelligence Tests

Specific reading disability, as with all other learning disabilities, cannot be diagnosed with a single test; it is a clinical diagnosis that requires a history of core symptoms, such as poor phonological awareness, inaccurate and slow oral reading, and poor spelling (Shaywitz. and Shaywitz 2013). While the SRD evaluation process requires the consideration of information from a variety of sources, including academic records, achievement testing, and classroom observation, cognitive testing remains fundamental to informed and accurate SRD identification. Cognitive testing is essential to differential diagnosis, distinguishing among multiple possible causes of reading underachievement. Furthermore, when used by skillful clinicians, cognitive testing is not only useful and effective for diagnostic purposes but provides information about individual needs that can be used to inform educational decisions and guide interventions (Decker et al. 2013).

Sometimes critics of cognitive ability testing claim that the results from intelligence tests are not useful for instructional planning. Often, however, the findings from these tests lead directly to specific alterations in a program. As an example, it would be very important for a teacher to know that a student had high verbal abilities but low reading skills. For this student oral evaluations would be the least biased, and the most accurate, way of assessing his or her actual knowledge and progress. If, on the other hand, a student had low verbal abilities as well as reading skills, he or she will likely need a modification in the difficulty level of evaluations and reading assignments. Another student with adequate verbal abilities and good reasoning may have a significant weakness in memory. This student may need specific compensatory aids to address this weakness, such as open book/open note testing and/or take-home exams. Another student with good memory but slow processing speed may have difficulty with completing clerical operations and more detailed procedures, such as note taking. An important accommodation for this student may be a note taker or a complete set of instructor notes.

Historically, school psychologists have over-relied on full scale intelligence test scores and arbitrary cutoff points for making special education eligibility decisions. This approach lacks psychometric and social validity and has consequently fallen out of favor in recent years.

Likewise, the future of intelligence testing is unlikely to involve reliance on the ability-achievement discrepancy procedure for SRD identification. While this procedure may have some application, as with the identification of twice exceptional individuals, it is neither sufficiently reliable nor valid to warrant continued widespread use.

While some have looked to RTI to replace the ability-achievement discrepancy procedure in the identification of children with SRD, the RTI process is not diagnostic and does not differentiate among the various causes of reading underachievement (Mather and Kaufman 2006). Furthermore, it provides little educationally useful information concerning the specific cognitive weakness underlying any specific case of SRD. Consequently, cognitive testing is likely to remain an important tool in the identification and diagnosis of specific reading disability for the foreseeable future. In the absence of these tests, it is difficult to distinguish between the diverse causes of reading disability or to select interventions that are targeted to the specific strengths and weaknesses of individual students.

In the future, the primary role of intelligence testing will likely lie in the focused use of cognitive assessment to detect reading disabilities (Fiorello and Primerano 2005) with the aim of understanding the child and identifying specific cognitive strengths and weaknesses (Kaufman 2006). The information derived from these assessments will be used both for identification and diagnostic purposes, as well as for informing educational decisions and guiding interventions. In fact, an understanding of an individual's comprehensive cognitive profile is key to interpretation and effective remediation (Kaufman et al. 2005). As observed by Kaufman (2006): "They aren't intelligence tests anymore; they are measures of cognitive processing." This observation not only predicts the future of cognitive testing but also takes us back to the past, prior to the use of an ability-achievement discrepancy, when the major purpose of cognitive assessments was to increase understanding of an individual's presenting problem, as well as his or her specific abilities. In 1937, Stanger and Donohue explained: "If these tests will give us a basis from which we can start to understand a child's difficulties, they will have justified the time spent on them. Anything that helps educators or parents to *understand* any phase of development or lack of development is of immeasurable value" (p. 189). More recently, Carroll (1993) observed that for several thousand years, "it has been recognized that there are individual differences in cognitive abilities, and that these differences have something to do with the roles and behaviors of individuals in society" (Carroll 1993, p. 25). Clearly, the future use of intelligence tests lies in increasing our understanding of how various cognitive processes predict and contribute to SRD, so students can receive timely, individualized, interventions.

Acknowledgments We would like to express our gratitude to Dr. Michael E. Gerner and Dr. Tobias Schneider-Richardson for their helpful comments.

References

Abu-Hamour, B., Urso, A., & Mather, N. (2012). The relationships among cognitive correlates and irregular word, non-word, and word reading. *International Journal of Special Education, 27*(1), 144–159.

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

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Publishing.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington: American Psychiatric Publishing.

Anastasi, A. (1980). Abilities and the measurement of achievement. In W. B. Schrader (Ed.), *Measuring achievement: Progress over a decade: Proceedings of the 1979 ETS invitational conference* (pp. 1–10). San Francisco: Jossey Bass.

Berninger, V. W. (1996). *Reading and writing acquisition: A developmental neuropsychological perspective*. Boulder: Westview Press.

Berninger, V., & Swanson, H. L. (2013). Diagnosing and treating specific learning disabilities in reference to the brain's working memory system. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 169–185). New York: Guilford Press.

Bijou, S. W. (1942). The psychometric pattern approach as an aid to clinical analysis—a review. *American Journal of Mental Deficiency, 46*, 354–862.

Boden, C., & Kirby, J. R. (1995). Successive processing, phonological coding and the remediation of reading. *Journal of Cognitive Education, 4*, 19–31.

Bowers, P. G., & Wolf, M. (1993). Theoretical links between naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing: An Interdisciplinary Journal, 5*, 69–85.

British Dyslexlia Association. (2009). *Dyslexia research information*. Retrieved from, http://www.bdadyslexia.org.uk/about-dyslexia/further-information/dyslexia-research-information-.html

Carlson, J., & Das, J. P. (1997). A process approach to remediating word decoding deficiencies in Chapter 1 children. *Learning Disabilities Quarterly, 20*, 93–102.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, MA: Cambridge University Press.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38*, 592.

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36*(2), 151–164.

Catts, H. W., & Kamhi, A. G. (Eds.). (2005). *The connections between language and reading disabilities*. Mahwah: Lawrence Erlbaum.

Crepeau-Hobson, F., & Bianco, M. (2011). Identification of gifted students with learning disabilities in a response-to-intervention era. *Psychology in the Schools, 48*, 102–109.

Crews, K., & D'Amato, R. C. (2009). Subtyping children's reading disabilities using a comprehensive neuropsychological measure. *International Journal of Neuroscience, 119*, 1615–1639.

Das, J. P. (1999). *PASS Reading enhancement program*. Deal: Sarka Educational Resources.

Das, J. P. (2002). A better look at intelligence. *Current Directions in Psychological Science, 11*(1), 28–33.

Das, J. P. (2009). *Reading difficulties and dyslexia: An interpretation for teachers*. New Delhi: SAGE International Publishers.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Needham Heights: Allyn & Bacon.

Das, J. P., Mishra, R. K., & Pool, J. E. (1995). An experiment on cognitive remediation or word-reading difficulty. *Journal of Learning Disabilities, 28*, 66–79.

Das, J. P., Parrila, R., & Papadopoulos, R. (2000). Cognitive education and reading disability. In A. Kozulin & Y. Rand (Eds.), *Experiences of mediated learning* (pp. 274–291). Amsterdam: Pergamon Press.

Decker, S. L., Hale, J. B., & Flanagan, D. P. (2013). Professional practice issues in the assessment of cognitive functioning for educational applications. *Psychology in the Schools, 50*, 300–313.

Ehri, L., Nunes, S., Willows, D., Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*, 250–287.

Evans, J. J., Floyd, R. G., McGrew, K. S., & Leforgee, M. H. (2002). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review, 31*, 246–262.

Feifer, S. G. (2008). Integrating response to intervention (RTI) with neuropsychology: A scientific approach to reading. *Psychology in the Schools, 45*, 812–825.

Feldmann, G., Kelly, R., & Diehl, V. (2004). An interpretative analysis of five commonly used processing speed measures. *Journal of Psychoeducational Assessment, 22*, 151–163.

Ferrer, E., Shaywitz, B., Holahan, J., Marchione, K., & Shaywitz, S. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science, 21*, 93–101.

Fiorello, C. A., & Primerano, D. (2005). Research into practice: Cattell-Horn-Carroll cognitive assessment in practice: Eligibility and program development issues. *Psychology in the Schools, 42*, 525–536.

Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1992). The validity of discrepancy-based definitions of reading disabilities. *Journal of Learning Disabilities, 25*, 555–561.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., … & Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology*, *86*, 6–23.

Fuchs, D., Fuchs, L. S., & Compton, D. (2012). Smart RTI: A next-generation approach to multilevel prevention. *Exceptional Children, 78*, 263–279.

Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences, 21*(1), 85–95.

Gallagher, J. J. (1966). Children with developmental imbalances: A psychoeducational definition. In W. M. Cruickshank (Ed.), *The teacher of brain-injured children* (pp. 23–43). New York: Syracuse University Press.

Georgiou, G. K., & Parrila, R. (2013). Rapid automatized naming and reading. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 169–185). New York: Guilford Press.

Gregg, N., Coleman, C., Davis, M., Lindstrom, W., & Hartwig, J. (2006). Critical issues for the diagnosis of learning disabilities in the adult population. *Psychology in the Schools, 43*, 889–899.

Haddad, F. A., Garcia, Y. E., Naglieri, J. A., Grimditch, M., McAndrews, A., & Eubanks, J. (2003). Planning facilitation and reading comprehension: Instructional relevance of the PASS theory. *Journal of Psychoeducational Assessment, 21*, 282–289.

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. New York: Guilford Press.

Hallahan, D. P., & Mercer, C. D. (2002). Learning disabilities: Historical perspectives. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 1–67). Mahwah: Lawrence Erlbaum Associates.
Hinshelwood, J. (1917). *Congenital word-blindness*. London: Lewis.
Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 267–300). New York: Wiley.
Horn, J. L., & Blankson, A. N. (2012). Foundations for better understanding of cognitive abilities. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 73–98). New York: Guilford.
Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). New York: Guilford Press.
International Dyslexia Association (2007). Frequently asked questions: What is dyslexia? Retrieved from http://www.interdys.org/FAQWhatIs.htm.
Iseman, J. S., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD and LD: A randomized controlled study. *Journal of Learning Disabilities, 44*, 184–195.
Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology, 17*, 209–224.
Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (Eds.). (2007). *Handbook of response to intervention*. London: Springer.
Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109*, 490–501.
Kail, R., & Hall, L. K. (1994). Processing speed, naming speed, and reading. *Developmental Psychology, 30*, 949–954.
Kail, R., Hall, L. K., & Caskey, B. J. (1999). Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics, 20*, 303–314.
Kaufman, A. S. (2006). *From David Wechsler to the new IDEA guidelines: 35 years in the eye of the IQ storm*. Anaheim: National Association of School Psychologists Annual Convention.
Kaufman, A. S. (2008). Neuropsychology and specific learning disabilities: Lessons from the past as a guide to present controversies and future clinical practice. In E. Fletcher-Janzen & C. R. Reynolds (Eds.), *Neuropsychological perspectives on learning disabilities in the era of RTI: Recommendations for diagnosis and intervention* (pp. 1–13). Hoboken: Wiley.
Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II Assessment*. New York: Wiley.
Kavale, K.A. (2002). *Discrepancy models in the identification of learning disability*. In R. Bradley, L. Danielson, & D.P. Hallahan (Eds.), Identification of learning disabilities: Research to practice (pp. 369–426). Mahwah, NJ: Erlbaum.
Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wechsler Scale profiles and recategorizations: Patterns or parodies? *Learning Disability Quarterly, 7*, 136–156.
Kavale, K. A., & Reese, J. H. (1992). The character of learning disabilities: An Iowa profile. *Learning Disability Quarterly, 15*, 74–94.
Kavale, K. A., Fuchs, D., & Scruggs, T. E. (1994). Setting the record straight on learning disabilities and low achievement. *Learning Disabilities Research & Practice, 9*, 70–77.
Kavale, K. A., Holdnack, J. A., & Mostert, M. P. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28*, 2–16.
Klassen, R. M., Neufeld, P., & Munro, F. (2005). When IQ is irrelevant to the definition of learning disabilities. *School Psychology International, 26*, 297–316.
Konold, T. R., Juel, C., McKinnon, M., & Deffes, R. (2003). A multivariate model of early reading acquisition. *Applied Psycholinguistics, 24*, 89–112.
Kussmaul, A. (1877). Disturbances of speech (trans: McCreery, J. A.). In H. von Ziemssen (Ed.), *Cyclopedia of the practice of medicine* (p. 595). New York: William Wood.
Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science, 20*, 1040–1048.
Lohman, D. F. (2006). Beliefs about differences between ability and accomplishment: From folk theories to cognitive science. *Roeper Review, 29*, 32–40.
Luria, A. R. (1966). Human brain and psychological processes. New York: Harper & Row.
Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading: A definition of dyslexia. *Annals of Dyslexia, 53*, 1–14.
Mather, N., & Gerner, M. (2009). Postsecondary students with high abilities and reading disabilities: Case analyses and commentary. *Learning Disabilities: A Multidisciplinary Journal, 15*, 121–130.
Mather, N., & Kaufman, N. L. (2006). Introduction to the special issue, part one: It's about the what, how well, and the why. *Psychology in the Schools, 43*, 747–752. doi:10.1002/pits.20185.
Mather, N., Shaywitz, B. A., & Shaywitz, S. E. (2013). Toward a synthesis of cognitive-psychological, medical-neurobiological, and educational models for the diagnosis and management of dyslexia. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The Oxford handbook of child psychological assessment* (pp. 698–721). New York: Oxford University Press.

McGrath, L. M., Pennington, B. F., Shanahan, M. A., Santerre-Lemmon, L. E., Barnard, H. D., Willcutt, E. G., … & Olson, R. K. (2011). A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: Searching for shared cognitive deficits. *Journal of Child Psychology and Psychiatry*, *52*, 547–557.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136–182). New York: Guilford.

McGrew, K., & Flanagan, D. (1998). *The intelligence test desk reference: Gf-Gc cross-battery assessment*. Needham Heights: Allyn & Bacon.

McGrew, K. S., & Wendling, B. J. (2010). Cattell–Horn–Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools, 47*, 651–675.

Monroe, M. (1932). *Children who cannot read*. Chicago: University of Chicago Press.

Monroe, M., & Backus, B. (1937). *Remedial reading*. Boston: Houghton Mifflin.

Morgan, W. P. (1896). Word blindness. *British Medical Journal, 1378*, 98.

Naglieri, J. A. (1999). How valid is the PASS theory and CAS? *School Psychology Review, 28*, 145–162.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive Assessment System. Administration and scoring manual. Interpretive handbook*. Austin: PRO-ED.

Naglieri, J. A., & Gottling, S. H. (1995). A study of planning and mathematics instruction for students with learning disabilities. *Psychological Reports, 76*, 1343–1354.

Naglieri, J. A., & Gottling, S. H. (1997). The PASS theory and mathematics instruction: A summary of initial studies. *Journal of Cognitive Education, 5*, 209–216.

Naglieri, J. A., & Johnson, D. (2000). Effectiveness of a cognitive strategy intervention to improve math calculation based on the PASS theory. *Journal of Learning Disabilities, 33*, 591–597.

Naglieri, J. A., & Conway, C. (2009). The Cognitive Assessment System. In J. A.

Naglieri, J. A., & S. Goldstein (2009). A practitioner's guide to assessment of intelligence and achievement (pp. 3–10). New York: Wiley.

Naglieri, J. A., & Pickering, E. B. (2010). *Helping children learn: Intervention handouts for use at school and home* (2nd ed.). Baltimore: Paul H. Brookes Publishing.

Naglieri, J. A., Salter, C. J., & Edwards, G. H. (2004). Assessment of children with attention and reading difficulties using the PASS theory and Cognitive Assessment System. *Journal of Psychoeducational Assessment, 22*, 93–105.

Naglieri, J. A. & Otero, T. (2011). Cognitive Assessment System: Redefining intelligence from a neuropsychological perspective. In A. Davis (Ed.). *Handbook of pediatric neuropsychology* (pp. 320–333). New York: Springer Publishing.

Naglieri, J. A., Das, J. P. & Goldstein, S. (2014). *Cognitive Assessment System* – Second Edition. Austin: PRO-ED.

O'Donnell, L. E. (1980). Intra-individual discrepancy in diagnosing specific learning disabilities. *Learning Disability Quarterly, 3*, 10–18.

Orton, S. T. (1925). Word-blindness in school children. *Archives of Neurology and Psychiatry, 14*, 581–615.

Parrila, R. K., Das, J. P., Kendrick, M., Papadopoulos, T., & Kirby, J. (1999). Efficacy of a cognitive reading remediation program for at-risk children in Grade 1. *Developmental Disabilities Bulletin, 27*, 1–31.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101*, 385–413.

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet, 379*(9830), 1997–2007.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology, 98*(1), 46–68.

Proctor, B., & Prevatt, F. (2003). Agreement among four models used for diagnosing learning disabilities. *Journal of Learning Disabilities, 36*, 459–466.

Restori, A. F., Katz, G. S., & Lee, H. B. (2009). A critique of the IQ/Achievement discrepancy model for identifying specific learning disabilities. *Europe's Journal of Psychology, 5*(4), 128–145.

Schultz, E. K., Simpson, C. G., & Lynch, S. (2012). Specific learning disability identification: What constitutes a pattern of strengths and weaknesses? *Learning Disabilities: A Multidisciplinary Journal, 18*, 87–97.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV Tests of Cognitive Abilities*. Rolling Meadows, IL: Riverside Publishing.

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., & DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*, 585–602.

Shaywitz, S. E., & Shaywitz, B. A. (2013). Making a hidden disability visible: What has been learned from the neurobiological studies of dyslexia. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 643–657). New York: Guilford Press.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22*, 469–486.

Siegel, L. S., & Mazabel, S. (2013). Basic cognitive processes and reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 186–213). New York: Guilford Press.

Snowling, M. J., & Hulme, C. (2012). Annual research review: The nature and classification of reading disorders—a commentary for proposals on DSM-5. *Journal of Child Psychology and Psychiatry, 53*, 593–607.

Speece, D. (2002). Classification of learning disabilities: Convergence, expansion, and caution. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 467–519). Mahwah: Erlbaum.

Stanger, M. A., & Donohue, E. K. (1937). *Prediction and prevention of reading difficulties*. New York: Oxford University Press.

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39*, 469–518.

Stuebing, K. K., Fletcher, J. M., Branum-Martin, L., & Francis, D. J. (2012). Evaluation of the technical adequacy of three methods for identifying specific learning disabilities based on cognitive discrepancies. *School Psychology Review, 41*, 3–22.

Swanson, H. L. (2013). Meta-analysis of research on children with learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 627–642). New York: Guilford Press.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research, 68*(3), 277–321.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39*, 252–269.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading. A meta analysis of the correlational evidence. *Review of Educational Research, 73*, 407–444.

Swanson, H. L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities, 42*, 260–287.

Travis, L. E. (1935). Intellectual factors. In G. M. Whipple (Ed.), *The thirty-fourth yearbook of the National Society for the Study of Education: Educational diagnosis* (pp. 37–47). Bloomington: Public School Publishing Company.

U. S. Department of Education. (2004). *Individuals with disabilities education improvement act, 20 U.S.C. § 1400*. Washington, DC: Author.

U.S. Office of Education. (1976). Proposed rulemaking. *Federal Register, 41*(230), 52404–52407.

Vaessen, A., & Blomert, L. (2013). The cognitive linkage and divergence of spelling and reading development. *Scientific Studies of Reading, 17*, 89–107.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*, 137–146.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A., & Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110*(1), 157–172.

Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology, 27*, 35–78.

Willcutt, E. G., Betjemann, R. S., McGrath, L. M., Chhabildas, N. A., Olson, R. K., DeFries, J. C., & Pennington, B. F. (2010). Etiology and neuropsychology of comorbidity between RD and ADHD: The case for multiple-deficit models. *Cortex, 46*, 1345–1361.

Willis, J. O., & Dumont, R. P. (2002). Guide to identification of learning disabilities (3rd ed.). Peterborough, NH: Authors. Available from authors: print copy from johnzerowillis@yahoo.com or CD from dumont @ fdu.edu.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*, 415–438.

27 Executive Functioning and Intelligence

Emily C. Duggan and Mauricio A. Garcia-Barrera

Over the last few decades, the conceptualization of the most complex cognitive processes of the human brain has received tremendous attention from researchers. As a result, cognitive models and theories of higher-level functioning are becoming increasingly detailed, sophisticated, and evidence-based. The proliferation of research in these areas, particularly intelligence and executive functioning, is driven by several key factors, including progress in neuroscience research approaches and the ever-increasing interest among researchers, clinicians, and policy makers to understand the social importance and implications of these cognitive abilities. Though we have made exceptional progress in our efforts to understand their principal neural correlates and underlying neurophysiologic mechanisms, intelligence and executive functioning are still nebulous—in all regards—and this in itself is the crux of the problem.

Interest in studying higher cortical processing, its relation to a specific brain area, and its impairments after neurological damage can be traced back to the Egyptians (circa 3500 BC) and their early yet sophisticated studies about human behavior and illness. However, the development of a theory of frontal lobe relations to human "higher" cognitive behavior is more recent. Early descriptions of this neuroanatomical and functional relationship include Harlow's famous case of Phineas Gage and the passage of an iron bar through his head, originally presented in 1848 as a letter to the editor of the *Boston Medical and Surgical Journal* and recently included in *The Journal of Neuropsychiatry and Clinical Neurosciences* (Harlow 1848 as reproduced by Neylan 1999). Mr. Gage was severely injured on his left, and probably right, prefrontal areas, but was able to walk and talk immediately after the accident, failing to demonstrate signs of behavioral change in the eyes of Harvard's Medical School examiners. Twenty years later, Harlow (1868) published a compilation of the behavioral changes observed in Mr. Gage during the 12 years he lived after the injury. The list included impulsivity, impatience, and irreverence, among other symptoms that interfered with Gage's capacity to perform at work and in personal relationships. These historical notes constituted the earliest description of what became known as "a frontal lobe syndrome." Interestingly, the early reports on Gage's recovery appear to identify dissociation in the presentation of his symptoms; that is, while Gage's behavior was childish and immature, his overall intellectual capacities and ability to independently navigate the demands of the environment seemed to have been mostly spared from damage. In fact, a characteristic of

E.C. Duggan, M.Sc. (✉) • M.A. Garcia-Barrera, Ph.D.
Department of Psychology, University of Victoria,
PO Box 1700 STN CSC, V8W 2Y2
Victoria, BC, Canada
e-mail: eduggan@uvic.ca; mgarcia@uvic.ca

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_27,

the frontal lobe syndrome post brain injury has been a relatively intact intellectual capacity with respect to premorbid levels but a marked impairment of executive functioning. This dissociation lies at the foundation of this chapter. Are executive functions associated with intelligence? What are the possible relationships between these two constructs? Given our recent progress in teasing out and characterizing cognitive functions, is there enough evidence of dissociation between these two complex umbrella constructs?

To address these questions, we will discuss the definitions of intelligence (briefly) and executive functions (to a larger extent), followed by a review of the empirical evidence provided by psychometric and neurophysiologic approaches to the examination of these two constructs' overlap, unique features, and their interactions.

Defining and Conceptualizing Intelligence: A Brief Review

In addressing the concept of intelligence, Robert Sternberg once wrote that when "looked at in one way, everyone knows what intelligence is; looked at in another way, no one does" (Sternberg 2000, p. 3). Traditionally, theories of intelligence have been founded on the existence of a general factor that contributes to performance on diverse cognitive tests. This factor was labeled general intelligence, or *g* (Spearman 1927). Though *g* has served great psychometric utility, its definition does little to practically capture the essence of *actual* intelligence. Additionally, the study of intelligence has a rich history and definitions of this construct are numerous and diverse. Despite the absence of a universal definition of intelligence, there are key elements common to nearly all of its definitions and conceptualizations. Thus, instead of rigorously defining intelligence, which is the purpose of previous chapters in this book, let us briefly explore the core essential aspects of intelligence that would inform our discussion about the relationship between this construct and that of executive function.

Two symposia, one in 1921 and a second in 1986, asking a panel of experts to define intelligence demonstrated that even diverse conceptions of intelligence share fundamental attributes (Sternberg and Detterman 1986; Terman 1921). According to these symposia, intelligence—at a minimum—involves basic mental processes, adaptation to the environment, and the ability to learn from experience. This definition has broadened over time to emphasize the importance of metacognition and higher-order thinking (e.g., decision making, problem solving, and reasoning). In a survey of 1,020 experts, essential elements of intelligence were identified based on the percentage of respondents rating them as important. These core elements with the highest percentages included: abstract thinking or reasoning (99.3 %), problem solving ability (97.7 %), capacity to acquire knowledge (96.0 %), memory (80.5 %), and adaptation to one's environment (77.2 %) (Snyderman and Rothman 1987).

Additional definitions of intelligence provided by prominent researchers over nearly 100 years include similar themes. Binet and Simon (1916) regarded intelligence as comprising of three distinct components: *direction* (identifying what needs to be done and how to do it), *adaptation* (customizing an environment-appropriate strategy for executing a task, then monitoring and adjusting the strategy through its implementation), and *criticism* (critiquing one's own thoughts and actions). In Spearman's two-factor theory of intelligence, the "general intelligence," or *g*, factor pervades all intellectual performances and accounts for the general mental energy that is involved in the most complex mental activities such as deductive operations (Spearman 1904, 1927). Intelligence has been viewed as the ability to accomplish "abstract thinking" (Terman 1921) and to undertake activities delineated not only by abstractness but also by difficulty, complexity, economy, adaptiveness to a goal, and social value (Stoddard 1943, p. 4). Similarly, it has been defined as "the capacity to reorganize one's behavior patterns so as to act more effectively and more appropriately in novel situations […] and the ability to carry on abstract thinking […] and problem solve" (Freeman 1955, pp. 149–50). Wechsler (1958) described 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" (p. 7), and Das (1973) defined it as "the ability to plan and structure one's behavior with an end in view" (p. 27).

One of the most influential taxonomies of the definition of intelligence was proposed by Horn and Cattell (1967). Their model included two types of intelligence: *fluid intelligence* (i.e., the ability to think logically and solve problems in novel situations) and *crystallized intelligence* (i.e., the ability to use knowledge, skills, and experience; Cattell 1963; Horn and Cattell 1967). Detterman (1986) commented, "intelligence can best be defined as a finite set of independent abilities operating as a complex system" (p. 57), while Estes (1986) asserted, "intelligence […] is a multifaceted aspect of the processes that enable animate or inanimate systems to accomplish tasks that involve information processing, problem solving, and creativity" (p. 66). Gardner (1983) wrote that intelligence "must entail a set of skills of problem solving—enabling the individual to resolve genuine problems or difficulties" (pp. 60–1). According to Campione and Brown (1978), intelligence has two basic components: an architectural (structural) system—comprised of capacity, durability, and efficiency—and an executive (control) system—constituting one's knowledge base, schemes, control processes, and metacognition. Das et al. (1994) describe cognitive ability as a function of planning (i.e., cognitive control, intentionality, knowledge, and self-regulation), attention, simultaneous processing (i.e., perception of the whole), and successive processing (i.e., decision making). Finally, an American Psychological Association consensus panel stated that with regards to intelligence, "individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" (Neisser et al. 1996, p. 77).

The considerable sampling of intelligence definitions summarized here illustrates the generally consistent and longstanding consensus among researchers despite their eclectic range of (often dissonant) orientations and perspectives—be it psychometric, cognitive, social, developmental, educational, and so on. Overall, we can surmise to say that *intelligence constitutes a confluence of the abilities to understand complex ideas and use experience and reasoning to solve problems and adapt to the environment*. Based on this definition, we can reasonably appreciate that intelligence represents the integration of numerous cognitive functions including attention, perception, working memory, abstract representation, planning, problem solving, memory, learning, and language. In fact, the integrative nature of intelligence has become a key area of focus within intelligence research (Colom et al. 2010; Roca et al. 2010) and may serve to help differentiate intelligence from other higher-level cognitive functions, as discussed below.

It is worth noting that in the last decades, progress toward elucidating the underlying neural physiology and mechanisms of intellectual functioning has been made. Intelligence has been correlated with greater total brain volumes (Wickett et al. 2000; McDaniel 2005), frontal, parietal, temporal, hippocampal, and cerebellar volumes (Luders et al. 2009; Toga and Thompson 2005), as well as total grey and white matter volumes, with a particular emphasis on the latter (Haier 2004, 2005; Gignac et al. 2003). A review of 37 structural and functional neuroimaging studies provided substantial evidence for a frontoparietal intelligence brain network that is highly distributed, with dorsolateral prefrontal cortex (Broca's areas 9, 45, 46, 47) and parietal cortex (BAs 7, 40) and the integrity of this network's white matter connectivity—particularly the arcuate fasciculus—being perhaps the most important for human intelligence (the Parieto-Frontal Integration Theory, Jung and Haier 2007; see also Chiang et al. 2009; Colom et al. 2009; Colom and Thompson 2011; Gläscher et al. 2009, 2010; Haier 2011). The integrity of the frontoparietal network and its harmonious coactivation is considered critical for successful intellectual functioning. Developmentally, general intelligence follows an inverted U-shaped trajectory (Jung and Haier 2007), though specific aspects of intelligence have been found to have unique trajectories. For example, fluid intelligence declines from the height of neurocognitive maturation around the mid-twenties, while crystallized intelligence continues to accumulate over the course of most of the life span—differences which result from a dynamic set of influences (Craik and Bialystok 2006).

More specific findings regarding the underlying mechanisms of intelligence depend on the way intelligence is defined and assessed. In the contemporary literature, intelligence is most typically framed in terms of psychometric *g*, fluid intelligence, and/or crystallized intelligence (Cattell 1963). As *g* is a higher-order, latent characteristic (and a purely statistical concept) of a complex cognitive system (Detterman 2000), it is typically measured with tests addressing numerous cognitive domains (e.g., Intelligence Quotient, or "IQ"). Interestingly, although most researchers would agree that fluid and crystallized intelligence are important but not exclusive components to the ecological manifestation of intelligence, often, fluid intelligence tasks of novel problem solving (e.g., Raven's Progressive Matrices; Raven 1938; Raven et al. 1988) or crystallized intelligence measures (e.g., vocabulary) are used in lieu of a full-scale IQ, as these tests have been shown to be excellent predictors of *g*. Advocacy for the latent variables approach (see *Evidence through psychometric studies* below) has more recently encouraged the use of diverse measures in the examination of higher-order cognition (Haier 2011; e.g., Martinez et al. 2013); however, skewness toward single-measure instruments for fluid intelligence has been a trend in research studies.

Using the basic framework that intelligence is a conceptual integration of our ability to understand complex ideas, reason, and draw upon experience to solve problems and adapt to the environment, let us now turn to another seemingly related higher-order cognitive construct tied to the same frontoparietal network: executive function.

Executive Function: A Definition

First defined by Luria (1973) and then labeled by Lezak (1982), the concept of "executive function" has proved to be at the top of the list of the most elusive constructs in neuropsychology (Jurado and Rosselli 2007). Despite its controversial existence, efforts to better understand humans' ability to exert a self-regulatory, goal-oriented, conscious control over action have remained a relevant issue. Importantly, current concerns about executive function as a construct of interest often include: (a) lack of a consensus about the definition, (b) lack of agreement about its structural fractionation (components), and (c) difficulty to reliably measure it (Garcia-Barrera et al. 2011). For the purpose of this chapter, we will only discuss issues surrounding the conceptualization and structure of executive functions.

Conceptualizing Executive Functions: Issues and Alternatives

One of the difficulties that researchers encounter when performing studies on executive functions is in choosing how to define this construct. A product of the diversity of theories is the plethora of terms that have been associated with this complex cognitive function. Often used terms range from "executive function" to "executive functions" with a variety of alternatives in between, such as "executive functioning," "central executive," "executive control," and "self-regulation." The term "executive system" appears to be the most adequate term for a theoretical conceptualization that emphasizes a modular system, but its operationalization in terms of behavior is difficult. This construct has been associated with the names "metacognition," "metacognitive function," and "metacognitive ability," with an emphasis on its relationship to awareness and self-regulation of problem solving, decision making, and learning skills. Finally, it has been historically recognized as "frontal lobe function" or "frontal functioning," terms that have been persistently used in some disciplines and even in some instruments; however, such names appear localizationist in nature, deemphasizing one of the most important characteristics of executive systems, that is, an extraordinary connectivity and coordinated networking within the brain. Moreover, this label seems oblivious to the bulk of research on the combined roles of prefrontal and parietal cortices in executive functioning, raising the question whether it would be more accurate to call it "prefrontoparietal function" instead.

The term "executive function" seems to imply the idea of a single-unit system, or acceptance of its role as an umbrella term; in contrast, the term "executive functions" would have the opposite effect, emphasizing the possibility of a range of them. For the purposes of this chapter, the term *executive functions* seems to be more appropriate, given three major reasons: first, it emphasizes the idea of diversity while allowing room for conceptual unity as a characteristic of the system; second, the term stresses the functional dimension of this construct more than its morphological association, allowing for the idea of cybernetic control as one of its functions rather than its sole function; and third, its behavioral operationalization appears more plausible, which is the ultimate objective when this construct is compared to other performance-based and behavioral-based constructs, such as intelligence.

Now, what are these executive functions? In a review of the neuropsychological literature, Jurado and Roselli (2007) offered one of the most compelling and clinically applicable definitions of this construct. These authors state that:

> In a constantly changing environment, executive abilities allow us to shift our mind set quickly and adapt to diverse situations while at the same time inhibiting inappropriate behaviors. They enable us to create a plan, initiate its execution, and persevere on the task at hand until its completion. Executive functions mediate the ability to organize our thoughts in a goal directed way and are therefore essential for success in school and work situations, as well as everyday living. (p. 214)

Influenced by a confluence between clinical neuropsychology and cognitive neuroscience perspectives, one integrated (representing unity) and yet multidimensional (representing diversity) definition should propose that executive behaviors are the outcomes of the interactions between cognitive and emotional control processes (Garcia-Barrera et al. 2012, 2013), aimed at the production of volitional, purposeful, and efficient guided behavior (Lezak 1982). These executive interactions are mediated by cortico-cortical connections within the prefrontal cortex and between prefrontal cortex and parietal cortex, as well as other corticothalamic-striato-cerebellar circuits (Fuster 2000), which are involved in rule setting and the organization of internal representations (Miller and Cohen 2001). When defined as outcomes, executive functions are highly variable across subjects, although there is some evidence that individual variability may be explained at large by genetic influences (Friedman et al. 2008).

Operationalization of Executive Functions: Gains from the Latent Variables Approach

Significant advances in the conceptualization and operationalization of the construct of intelligence were made possible with the introduction of latent variable analysis approach. Exploratory factor analysis (i.e., principal components) on psychometric tests yielded a set of factors (latent in nature), which served for the identification of the general factor *g* and its many underlying broad abilities (McGrew 2009). Similarly, and due to the latent nature of executive functions, significant progress in their conceptualization and operationalization has been accomplished in the last decade since the introduction of latent variable analysis approach to elucidate the possible underlying components within its complex structure (Miyake et al. 2000). In fact, a considerable amount of research has been dedicated to examine the latent structure of executive abilities in preschoolers (e.g., Espy et al. 2011; Hughes et al. 2009; Wiebe et al. 2011), kindergarteners (e.g., Sadeh et al. 2012), school-aged children and adolescents (e.g., Garcia-Barrera et al. 2011, 2013; Gur et al. 2012; Huizinga et al. 2006; Lehto et al. 2003), young adults (e.g., Frazer 2012; Miyake et al. 2000; Friedman and Miyake 2004; Friedman et al. 2008), and middle-age and older adults (e.g., Adrover-Roig et al. 2012; Hull et al. 2008).

Three executive functions have gained increasing attention due to their more feasible operationalization (e.g., Miyake et al. 2000): inhibition of goal-irrelevant representations (*inhibiting*), flexible attentional shift toward goal-relevant representations (*shifting*), and fluid update of mental representations in working memory as the goal is planned and executed (*updating*). However, efforts have been made to isolate other executive components that may also serve a role in the multiple system interactions that result in everyday executive behavior, for instance, problem representation and emotional control (Frazer 2012; Garcia-Barrera et al. 2011, 2012). Problem representation refers to the identification of novelty, organization of incoming and internalized inputs, and use of strategic seriation to generate a set of temporally organized rules (or syntax) for purposeful goals of action. Emotional control is associated with valence and reward processing and specifically refers to the top-down regulation of the impact of reward-saliency in behavior and biasing valence evaluation toward goal-based behavioral action (Frazer 2012).

The latent variable approach requires optimal operationalization of the executive functions. Therefore, one of the secondary gains of modeling executive functions as latent variables has been a robust psychometric approach to examining their development, adding to the discussion about the unity versus the diversity of executive functions across the life span. Although using mostly cross-sectional design, some research studies support one or the other side of the debate, and a few suggest that executive functions may develop following a unity to diversity gradient, possibly characterized by a baseline unidimensional (Wiebe et al. 2011; Hughes et al. 2009) or two-dimensional system (Espy et al. 1999; Miller et al. 2012), which may fractionate into a range of three or more executive subsystems between the ages of six to eight (Brocki and Bohlin 2004; Garon et al. 2008; Lehto et al. 2003). The multidimensionality of this latent variable-based construct seems feasible in light of models of neural specialization and fractionation of higher-order cognitive functions during the earlier neurodevelopmental stages (Tsujimoto 2008).

Further, an inverted U developmental trajectory of this system, delayed with respect to the development of other cognitive systems, has also been documented (e.g., Craik and Bialystok 2006). This delayed cognitive maturation is supported by developmental trajectories observed in frontal and parietal association cortices (Gogtay et al. 2004), which are strongly associated with executive functions (Collette et al. 2006). Longitudinal examination using the latent approach has demonstrated that the development of executive functions is nonlinear and that different within-subject developmental trajectories can be observed across executive abilities even as early as between the ages of six to eleven (Garcia-Barrera et al. 2013). Some executive abilities reach full maturity and plateau in their development by age 15 (e.g., shifting, inhibition; Huizinga et al. 2006), whereas other executive abilities may continue unfolding well into young adulthood (e.g., working memory; Huizinga et al. 2006), reaching a plateau around the mid-20s (McAuley and White 2011). These differential developmental trajectories have been observed independently from the gains in speed of processing information associated with age, which, as it would be expected, count for a portion of the variance in test performance.

The multidimensionality of executive functions appears to remain somewhat stable during the earlier adulthood (McAuley and White 2011), but little is known about the patterns of differential decline through later adulthood, given inconsistent research findings (Jurado and Rosselli 2007). Some researchers have postulated a dedifferentiation hypothesis, that is, the idea that cognitive abilities (including executive skills) become less differentiated with age, as correlations between executive functions significantly increase with age, and groups of executive tasks are reducible to a single factor in older adults. In other words, this hypothesis postulates that through development, cognitive abilities differentiate from an initial unity in early development, and then return to dedifferentiation as adults age (Balinsky 1941; also discussed in Anstey et al. 2003; and illustrated in Fig. 27.1). However, some have argued that this dedifferentiation may be related to age-related changes in fluid intelligence more than executive functions per se (Salthouse 2001). It is known though that the patterns of cognitive decline associated with age follow closely the reduction in network connectivity and structural atrophy observed in the elderly (Craik and Bialystok 2006; Jurado and Rosselli 2007), keeping in mind the interactions between brain changes and cognition are nonlinear, dynamic, and significantly moderated by genetic, environmental, and social factors (Baltes 1987).

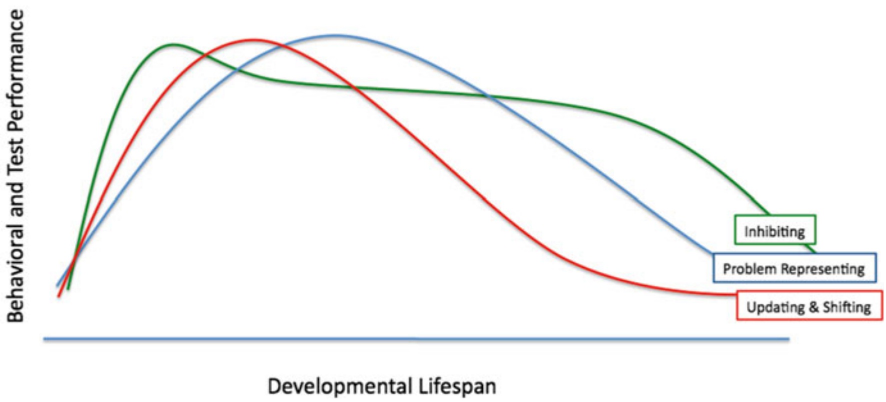


Fig. 27.1 Speculative model of the developmental trajectory of three executive functions. Emphasis is placed on the different rates of development during childhood and the differential decline with aging

Understanding the Relationship Between Executive Functioning and Intelligence

A glance into the aforementioned definitions of intelligence and executive functions is enough to identify the conceptual commonalities between these two constructs. When defined as the ability to problem solve, plan, and structure goal-directed behavior, to produce convergent logical reasoning in the face of novelty, and to make not only efficient but efficacious decisions, little is left for one to distinguish intelligence from executive functioning. The field of neuropsychological assessment has traditionally favored more general descriptors such as "general cognitive ability" or simply "cognitive" or "mental" abilities to refer to the group of processes that underlie the umbrella term "intelligence" (Lezak et al. 2012), avoiding part of the confusion.

Despite the conflation observed between these two constructs, a number of efforts have been made to examine possible divergent components and to better identify the commonalities. Progress has been made particularly through the use of the psychometric approach, the study of neural correlates, and the examination of clinical populations; discussion of these lines of inquiry follows.

Findings from Psychometric Studies

Historically, the construct of intelligence preceded that of executive function by a large time difference. Since the creation of instruments to assess intelligence began, neuropsychologists have been focused on the discussion about the limitations of interpreting the Intelligence Quotient (IQ) as a single composite to measure a range of cognitive abilities as diverse as attention, memory, reasoning, and processing speed. Thus, the unity versus diversity debate, often associated with executive functions (e.g., Miyake et al. 2000), had long before been associated with intelligence. Is intelligence best represented as a single general cognitive capacity that changes over time, or is it best conceived as the outcome of the interactions between several cognitive abilities? Two of the most robust evidenced-based premises favoring the idea of diversity over that of a single *g* factor have been that "brain function is too complex to be communicated by a single score" and that a summation score does not have a "predictably direct relationship to the size of brain lesions" (Lezak et al. 2012, p. 22). In other words, for neuropsychologists concerned about brain-behavior relationships, a single score suggests a single localization, or at least nodal and vulnerable to the effects of the damage extension. Lezak and colleagues (2012) also pointed out that the patterns of cognitive decline observed in typical and atypical aging are characterized by the different rates in which abilities change. Similar assertions can be made for the earlier development of these abilities (Blair 2006).

In contrast, a psychometric perspective, using a latent variable approach such as factor analysis techniques, has been traditionally successful at confirming a structure of intelligence that generally integrates Spearman's *g* factor at the top of the hierarchy. The so-called Cattell-Horn-Carroll (CHC) model of intelligence (Blair 2010) integrates the factor-analytical work of John Carroll (1993) with the two-factor (i.e., fluid and crystallized) model of intelligence proposed by Cattell and Horn (1967) and has become the landmark psychometric structure of intelligence (McGrew 2009) making it a robust and attractive approach.

As Blair (2010) pointed out, and despite the statistical robustness that the latent variable approach offers, the psychometric tradition does come with several limitations. One of them includes the dependency of factor analysis on test selection. Due to the need of including performance measures of mental and cognitive abilities, tests often involve types of abilities that are too academic and not very ecologically valid (e.g., related to everyday life behavior; Floyd et al. 2010). Another limitation involves the fact that the traditional psychometric instruments involved in the derivation of the *g* factor do not include emotional and social aspects of intelligence (Lezak et al. 2012) or other multiple types of intelligence (Gardner 1983).

Efforts in distinguishing the psychometric components involved in the structure of intelligence have produced interesting definitions of the two main core factors loading to *g*, that is, fluid and crystallized intelligence. In a review of theories, McGrew (2009) captured the definition of fluid intelligence (Gf) as:

> The use of deliberate and controlled mental operations to solve novel problems that cannot be performed automatically. Mental operations often include drawing inferences, concept formation, classification, generating and testing hypothesis, identifying relations, comprehending implications, problem solving, extrapolating, and transforming information. Inductive and deductive reasoning are generally considered the hallmark indicators of Gf. (p. 5)

Whereas crystallized intelligence (Gc) was defined as:

> … a person's breadth and depth of acquired knowledge of the language, information and concepts of a specific culture, and/or the application of this knowledge. Gc is primarily a store of verbal or language-based declarative (knowing what) and procedural (knowing how) knowledge acquired through the investment of other abilities during formal and informal educational and general life experiences. (p. 5)

It is worth noting from these definitions the influence of educational experience in crystallized intelligence (Jonson and Bouchard 2005) and—more important for our discussion in this chapter—the resemblance between the definition of fluid intelligence and executive functions. Thus, how do fluid intelligence and executive functions differ? This question and similar ones (e.g., Are they differentially involved in behavior? Does one construct subsume the other?) have been pondered for decades, yet have been rarely empirically investigated (Floyd et al. 2010). Since the recognition in neuropsychology of the term executive functions, a strong conceptual relationship between intelligence and executive functions has been largely discussed in the literature. Sternberg (1985) proposed that executive functions are common to all cognitive tasks and that psychometric *g* merely represents individual differences in executive functioning.

In fact, significant correlations between intelligence test scores and executive functioning measures have been demonstrated (Ardila et al. 2000; Arffa 2007; Brydges et al. 2012; Burgess et al. 1998; Dempster 1991). Floyd and colleagues (2010) examined the question "How do executive functions fit with the Cattell-Horn-Carroll model?" by conducting a detailed psychometric analysis of the relationships between a well-established and known executive functions battery (the Delis-Kaplan Executive Function System-D-KEFS, Delis et al. 2001) and the Woodcock-Johnson III Tests of Cognitive Abilities (WJ-III; Woodcock et al. 2001). Their analyses demonstrated that measures of both constructs are not easily distinguished, that is, all D-KEFS subtests not only correlated with intelligence subtests but also measured both broad ability factors and the general factor ability, as demonstrated in their exploratory and confirmatory factor analyses (Floyd et al. 2010). Further, a factor they label Executive Function (EF) had the second most robust loadings to the general factor *g* (0.85), surpassed only by their comprehension-knowledge factor (0.95), the latter being a somewhat cluttered factor that included several D-KEFS tasks such as the sorting task (tapping similar abilities to the Wisconsin Card Sorting Test), tasks for verbal reasoning and inference (20 questions test and word context), and tasks from the WJ-III such as the analysis-synthesis, verbal comprehension, and general information. The comprehension-knowledge factor lies closely between Carroll's crystallized and fluid intelligence factors. However, when detailed confirmatory factor analysis is applied, some researchers have reported that crystallized knowledge tasks are not as strongly associated with performance on measures of executive functioning (e.g., Espy et al. 1999; Krikorian and Bartok 1998; Pennington 1997). In her wisdom, Denckla (1996) reminds us that this complex overlap between psychometric *g* and executive functioning should not be reduced to the idea that executive functioning is simply "old wine in new bottles" (p. 268). Crinella and Yu (1999) examined animal models, frontal lobe lesion studies, and psychometric studies on clinical populations and concluded that executive functioning although relevant to problem solving neither equals nor explains *g* in its totality.

If the general factor *g* is not a psychometric equivalent to executive functioning, is fluid intelligence a considerable equivalent? In the tradition of Baddeley's central executive (Baddeley 1996), Pennington (e.g., Pennington et al. 1996) proposed that executive functioning is best defined as the outcome of working memory, which he believes is closely associated with fluid intelligence, leaving the rest of *g* (crystallized intelligence) to be regarded as non-executive in nature. Pennington's predictions about executive functions may have left out relevant elements of this complex system, but the relationship between working memory and fluid intelligence has been psychometrically demonstrated to be high (Ackerman et al. 2005; Kane et al. 2005). We know from the work of Carroll (1993) that fluid intelligence is closely associated with *g*, creating the equation EFs(*WM*) ~ Gf ~ *g*, that is, working memory as a proxy for executive functioning is closely associated with fluid intelligence which is closely associated with *g*.

To examine this issue using a latent variable approach, Friedman et al. (2006) investigated the associations between performances on three components of executive functioning (i.e., inhibiting, updating working memory, and shifting) in relation to performance on both fluid and crystallized intelligence measures, in healthy young adults. Their findings are consistent with Pennington's only to some extent. These authors found a robust relationship between their updating working memory factor and both types of intelligence. Their confirmatory analysis followed by structural equation modeling confirmed that when the correlations between the three executive functions are accounted for (as per in Miyake et al. 2000), the associations between inhibiting and shifting with intelligence drop significantly, demonstrating that it is through their shared variance with updating working memory that the association between executive functioning and intelligence operates. Even though these associations were high (loadings of 0.74 with fluid intelligence and 0.79 with crystallized intelligence) and accounted for a significant portion of the variance in intelligence measures (37–45 %), the authors recognized that there is more to intelligence than just executive functioning or at least than just inhibiting, shifting, and updating. Salthouse and colleagues (Kail and Salthouse 1994; Salthouse and Pink 2008) assert that what is common between executive functions tasks and intelligence is not working memory but speed of processing. Examples of these interactions are abundant in the recent literature. In Floyd et al. (2010), a processing speed factor accounted for a significant portion of the variance in two out of the four subtests comprising their executive function factor. Adrover-Roig and colleagues (2012) used a speed of processing factor as a moderator latent variable to effectively examine the structure of executive functions in older adults. Speed of processing may play a relevant role in the development of cognitive and mental abilities of all sorts, and it is an important variable to account for when examining performance in the tasks often used to test these models. However, the difficulty in isolating the pure influence of this foundational process during any cognitive task remains a challenge.

Despite the robustness of some of these psychometric findings, Blair (2010) wisely points out that:

> Although working memory capacity would appear to be a primary determinant of Gf, and perhaps by extension the factor itself, somewhat ironically, Gf and EF have been shown in clinical and developmental research to be distinct from the general factor, as well as from other aspects of intelligence (Blair 2006). The finding that Gf is so highly similar to g in the psychometric literature but so clearly distinct from it in others, including clinical neuropsychology and developmental disabilities, […] has been a central point in the study of intelligence […] and also of the neurobiology of intelligence. (p. 231)

Acknowledging Blair's (2010) assertion, we now turn to reviewing the literature examining commonalities and differences between executive functioning and intelligence at the level of their possible neural correlates and when examination of clinical populations is employed.

Examining the Neural Correlates of Executive Functioning and Intelligence

Despite the wide and often conflicting array of viewpoints on intelligence and executive functioning, a substantial body of research evidence indicates that both are multifaceted constructs representing the cognitive products of similar (but not identical) neural networks. Overall, lesion and neuroimaging studies have demonstrated the significance of prefrontal and parietal areas and their connections to higher-level cognition including intelligence (Haier 2011), executive functioning (Niendam et al. 2012), and working memory (Wager and Smith 2003). An extensive body of research has been dedicated to understanding the underlying neurobiology and corresponding brain networks of intelligence and similarly, but to a lesser extent, of executive functioning; however, studies actively examining the relationship between intelligence and executive functioning have been fewer. Though entire volumes could be dedicated to documenting these lines of research and their points of convergence and divergence, here we attempt to briefly review the most central points of this literature (with an emphasis on recent studies) to provide a basis for our following discussion of contemporary research dedicated to examining executive functioning and intelligence in tandem.

Executive neurologic substrates versus intelligence neurologic substrates. Brain lesion analysis, or lesion-symptom mapping, has served as a fundamental research methodology for gaining insights into brain-behavior relationships and localizing cognitive functions. As previously indicated, modern approaches to lesion analysis in relation to intelligence, executive functions, and frontal lobe functioning trace their history back to Harlow's famous case of Phineas Gage (1848, 1868), as well as to the broader work of numerous influential figures including Paul Broca (1848–1904), Karl Wernicke (1848–1905), Hughlings Jackson (1835–1911), Norman Geschwind (1926–1984), and Alexander Luria (1902–1977). Research has demonstrated recruitment of prefrontal networks for performance on both tests of general intelligence (Bishop et al. 2008; Duncan et al. 2000; Esposito et al. 1999; Prabhakaran et al. 1997) and executive functions (Duncan 2006; Duncan and Owen 2000; Miller 2000; Miller and Cohen 2001), and prefrontal cortex lesions have been shown to result in a wide array of cognitive deficits, including numerous aspects of executive and intellectual functioning, depending on the damage (cf. Fuster 1988; Goldman-Rakic 1987; Luria 1973; Roberts et al. 1998; Stuss and Benson 1984; Wise et al. 1996). Though it seems reasonable that the prefrontal cortex could serve as the shared fundamental neurologic substrate of intelligence and executive functions, this argument has mixed support, especially when it comes to localizing within the level of the prefrontal cortex.

It is relatively common for patients with frontal lobe lesions to demonstrate preserved psychometric intelligence (Ackerly 1937; Hebb 1945, 1949; Hebb and Penfield 1940; Teuber 1959; Weinstein and Teuber 1957; Milner 1963; Bar-On et al. 2003; Black 1976; Janowski et al. 1989), and frontal lobe damage has been associated with marked impairments in executive functioning in spite of grossly preserved or normal intellectual abilities, as measured through conventional intelligence and neuropsychological tests (Bechara et al. 1998; Damasio et al. 1990; Eslinger and Damasio 1985; Saver and Damasio 1991; Tranel et al. 2008). Although there is a considerable body of evidence demonstrating that the frontal lobes are required for human and animal models of executive functioning (e.g., Cummings and Benson 1990; Eslinger and Damasio 1985; Teuber 1964), frontal lobe function and executive functioning is not a direct "one-to-one relationship" (see Alvarez and Emory 2006 for a review), and tests commonly used to assess executive abilities traditionally associated with prefrontal regions do not always show clear regional specificity (e.g., Wisconsin Card Sorting Test and dorsolateral prefrontal cortex, Milner 1963; Rezai et al. 1993), but also other frontal and non-frontal regions (Anderson et al. 1991; Barcelo and Santome-Calleja 2000; Buchsbaum et al. 2005; Horner et al. 1996; Stuss et al. 1983).

Preserved psychometric intelligence (as measured with a conventional IQ) following frontal lobe damage was challenged by Duncan et al. (1995, 1996) who demonstrated intelligence impairments in relation to frontal lobe deficits using high *g*-loadings measures of fluid intelligence (e.g., Progressive matrices, Culture Fair), which they argue are better for detecting intelligence changes after brain damage. Such use of fluid intelligence measures has received support from imaging studies showing consistent activation patterns within frontal and parietal areas and highlighting the particular importance of the lateral prefrontal cortex, the supplementary motor area/anterior cingulate, the anterior insula/operculum, and the intraparietal sulcus (Bishop et al. 2008; Duncan et al. 2000; Duncan and Owen 2000; Gray et al. 2003; Prabhakaran et al. 1997; Tranel et al. 2008). Nevertheless, lesion size and/or distribution has been a point of criticism (among several) in this literature, with similar methods being used to report affected executive, but not intellectual, functioning after focal frontal lesions. Furthermore researchers counter that intelligence is not an exclusive manifestation of executive functioning; rather, executive functioning represents a critical but not the only component of cognitive information processing required for problem solving (Crinella and Yu 1999; Kane and Engle 2002). Although intelligence and executive functions share many similar neural correlates including prefrontal and parietal association areas, accumulating evidence indicates that some areas including ventral and anterior prefrontal cortices are likely more unique to executive functioning (Collette et al. 2006; Juardo and Rosselli 2007; Wager and Smith 2003).

As discussed above (see Sect. "Findings from Psychometric Studies"), subsequent neuroscience models of intelligence have typically fallen under the scope of one of two orientations (i.e., intelligence as a single general cognitive capacity changing over time versus intelligence as the outcome of a numerous integrated cognitive abilities), both of which are supported by neurophysiologic data. First, deriving from Spearman's (1904, 1927) theory, the prefrontal cortex has been proposed as a unified neural architecture of the general factor underlying a wide range of cognitive abilities, or *g* (e.g., Duncan et al. 2000; cf. Roca et al. 2010; Barbey et al. 2012). Second, deriving from Thomson's (1951) theory (as discussed in Bartholomew et al. 2009), tests of intelligence have been proposed to represent the sum of numerous separate cognitive operations functionally mediated by specialized regions of the brain, particularly the frontal and parietal cortices and the white matter association tracts that network these areas (e.g., Colom and Thompson 2011; Gläscher et al. 2009, 2010; Jung and Haier 2007). Review of both models, interestingly enough, demonstrates that each can account for the pattern of positive intercorrelations among cognitive tests and current methods (i.e., psychophysiological, neuroimaging, genetic) cannot distinguish between them (Bartholomew et al. 2009; Deary et al. 2010). A third plausible model of intelligence, proposed by Van Der Maas and colleagues (2006), posits that cognitive processes grow over time and mutually interact, which can account for the positive correlations between various cognitive tests; however, some aspects of the model have received criticism, and it remains to be seen how successful it will fair when tested in less theoretical applications (Bartholomew et al. 2009).

These models bear great significance in the executive functions versus intelligence forum, as they simultaneously implicate executive functioning in their positioning (based on the significant overlap between the conceptualizations of intelligence and executive functions definitions) and guide the conceptual and experimental frameworks for a preponderance of current executive-intelligence research. Additionally, most lesion studies relevant to this discussion have been flawed to some extent, which has limited our ability to effectively evaluate executive functioning and intelligence with respect to one another. Barbey et al. (2012b; see also Barbey et al. 2012a) summarize in their review:

> Of the neuropsychological patient studies that have examined the neural basis of general intelligence (Basso et al. 1973; Black 1976; Eslinger and Damasio 1985; Shallice and Burgess 1991; Bechara et al. 1994; Duncan et al. 1995, 1996; Burgess and Shallice 1996; Isingrini and Vazou 1997; Parkin and Java 1999; Blair and Cipolotti 2000; Kane and Engle 2002; Bugg et al. 2006; Gläscher et al. 2009, 2010; Roca et al. 2010; Tranel et al. 2008) and executive functioning (Ptito et al. 1995; D'Esposito and Postle 1999; Muller et al. 2002; D'Esposito et al. 2006; Baldo and Dronkers 2006; Volle et al. 2008; Tsuchida and Fellows 2009), all share one or more of the following features: diffuse (rather than focal) brain lesions, lack of comparison subjects carefully matched for pre- and post-injury performance measures, and exclusive use of general intelligence or executive function tests. (p. 1155)

It is under this challenging context of issues and discrepancies, as well as myriad factors not addressed here, that the most recent investigations into intelligence and executive functioning operate.

Brain-lesioned patients and healthy controls. Recently, Roca and colleagues (2010) examined the role of fluid intelligence in a variety of executive tasks in two experiments with patients with focal frontal lesions. In one experiment, patients with frontal lesions were significantly impaired on measures of intelligence (Culture Fair IQ) and executive functioning (Wisconsin Card Sort Test and Verbal Fluency), executive functioning was correlated with intelligence, and higher IQ was associated with better executive function task performance, with differences between patients and controls "entirely explained by *g*" (Roca et al. 2010, p. 243). In a second experiment, although executive functioning tasks (Ineco Frontal Screening, Hotel Task, Iowa Gambling Task, Faux Pas test, and Mind in the Eyes task) were correlated with IQ (Culture Fair), with better performance associated with higher IQ, executive functioning differences between patients and controls remained even after correcting for IQ differences, and their results suggested these executive deficits to be associated with greater lesion volume in the right anterior prefrontal cortex. The results of this study are intriguing in that it appears that unique contributions of executive functioning tasks assessing abilities or behaviors most different from those traditionally associated with intelligence are more likely to persist after controlling for IQ. Also, considering the likelihood that components of both intelligence and executive functions contribute to psychometric *g*, in so far as they are conceptualized and measured, we agree that it is beneficial to examine executive functioning after controlling for intelligence, but we also suggest examining intelligence when controlling for executive functioning (see the discussion on Dichter et al. 2006, below in *Clinical Disorders in Adults*, as well as the discussions on Martinez et al. 2013; Gläscher et al. 2012, for an alternative perspective).

Using voxel-based lesion-symptom mapping, Barbey et al. (2012) found measures of *g* (WAIS) and executive functioning (D-KEFS) were associated with circumscribed damage to left lateralized regions and interconnections of the frontal and parietal cortices. Although their results indicate that executive functioning and intelligence fundamentally depend on shared neural substrates of the frontoparietal network, they also found areas uniquely related to one but not the other, and vice versa. Specifically, the right superior and inferior parietal lobe and left inferior occipital gyrus related to *g*, while the left anterior frontal pole was related to executive function.

Furthermore, in a second study of patients with focal dorsolateral prefrontal cortex lesions using a similar methodology, Barbey et al. (2012) found the dorsolateral prefrontal cortex functionally supports mechanisms for intelligence (corroborating the existing neuroscience literature and particularly the findings of Woolgar et al. (2010), linking dlPFC to fluid intelligence), but not for the mechanisms of executive functioning (as measured with the D-KEFS). In discussing their findings, Barbey et al. (2012) suggest that "*g* and executive function draw on the combination of conceptual knowledge and executive processes, and that the communication between areas associated with these capacities is of critical importance" (p. 1160). It is not readily apparent to us if such a conclusion does more to clarify or to obscure the already grey relationship between executive functions and intelligence; however, the strength of these authors' approach clearly lies in their use of the latent variables, and future studies in this field would benefit from implementation of similar methodological procedures.

The use of a priori hypotheses has been considered another factor limiting the scope of neuroanatomical conclusions that can be made about higher-order cognitive abilities; two studies abandoning the a priori approach have provided interesting contributions to the debate at hand (Gläscher et al. 2012; Martinez et al. 2013). First, Martinez et al. (2013) recently conducted a study in healthy controls to "identify brain networks whose properties predicted higher-order cognitive function" with fiber-tracking techniques and network analysis (p. 608). These authors found that "in general, specific networks were better predictors for each different intelligence measure, but only one was shared by them all [...] Temporal to lateral prefrontal connections were relevant to explaining the variability in general intelligence, fluid intelligence, crystallized intelligence, and mental processing speed" (p. 611). Though their study is geared more toward the intelligence component of higher-order cognitive functioning, these findings still hold implication for executive networks, and their approach may be used in the future to examine executive functions and intelligence together.

Second, using voxel-based lesion-symptom mapping in a large sample of patients with focal brain lesions, Gläscher et al. (2012) characterized the regional specificity of two functions (statistically isolated from performance across multiple tasks) previously associated with prefrontal networks: cognitive control and value-based decision making. These authors' findings revealed two distinct functional-anatomical prefrontal networks. First, cognitive control (response inhibition, conflict monitoring, and switching) was associated with the anterior cingulate cortex and dorsolateral prefrontal cortex. Second, value-based decision making (valuation, decision making, and reward learning) was associated with the ventromedial, frontopolar, and orbitofrontal areas of the prefrontal cortex. Additionally, "cognitive control tasks shared a common performance factor related to set shifting that was linked to the rostral anterior cingulate cortex… [while] regions in the ventral PFC were required for decision-making" (p. 14681). Gläscher et al. conclude that their findings provide detailed lesion-based evidence for a remarkable level of functional-anatomical specificity in the prefrontal cortex.

While the research findings presented by Martinez et al. (2013) and Gläscher et al. (2012) each contribute further to the understanding of the neural substrates of executive functions and intelligence, it is their approaches that are perhaps even more interesting and which could lead to a potential research avenue that (somewhat) bypasses the construct entanglement issue that is inherent to intelligence and executive functions.

Evidence from clinical studies. Perhaps the difficultly in elucidating the nature of the relationship between executive functions and intelligence is more broadly related to the populations we are hoping to gain insights from. As highlighted above, some similarities in the relevant psychometric literature have been found distinctly different in clinical research. Thus, we aim to review clinically driven areas of research that provide a slightly different perspective on conceptualizing and exploring the executive-intelligence relationship.

In the context of developmental disorders, the relationship between intelligence and executive functioning has perhaps been most thoroughly investigated in individuals with attention deficit hyperactivity disorder (ADHD). Many theorize that symptoms of ADHD arise from executive dysfunction (e.g., self-regulation and behavioral control problems), and early lines of research purported that executive dysfunction in ADHD could be attributed to IQ differences, due to the high correlation with many traditional executive function measures. Executive impairments have since been shown to persist despite normal range psychometric intelligence in children with ADHD, as well as in children with other developmental disorders commonly linked to executive dysfunction (e.g., phenylketonuria and specific learning disabilities; Barkley 1997; Berlin 2003; Diamond et al. 1997; McLean and Hitch 1999; Stanovich et al. 1997; Swanson 1999).

An example of the examination of the relationship of executive functions to intelligence in developmental disorders has most recently been addressed by Osório and colleagues in individuals with Williams syndrome, a rare neurodevelopmental disorder characterized by a distinct cognitive profile of strengths and weaknesses (i.e., impairments in nonverbal domains and executive functioning in contrast to relative verbal strengths) along with a pattern of other features (Osório et al. 2012). These researchers examined the relationships between IQ (WISC FSIQ) and three executive function components selected based on the work of Miyake et al. (Friedman et al. 2006; Miyake et al. 2000): working memory (reverse digit span), shifting (trail making test, Wisconsin card sorting task), and inhibiting (Stroop task, go/no-go). When compared with typically developing controls, individuals with Williams syndrome demonstrated significantly poorer executive functioning; however, after controlling for IQ most differences were lost, leading the authors to theorize that general cognitive ability may account for group executive function differences (Osório et al. 2012). Due to the characteristic cognitive pattern in Williams syndrome, this line of research offers a potentially significant opportunity to more closely examine crystallized aspects of intelligence in relation to fluid intelligence and executive functions, a research topic not readily being addressed in the current literature. However, some limitations on this approach must be recognized. In particular, this study depended on an inaccurate interpretation of the Miyake et al. (2000) executive function components (e.g., examining broader working memory rather than updating) and selecting tasks that are best explained by multiple executive domains (e.g., Wisconsin Card Sorting Test). Future studies should include tasks that examine more carefully the individual executive components and that allow for consideration of relevant clinical factors (e.g., verbal/visual strengths or weaknesses).

Another intriguing context in which executive functioning and intelligence relationships may be considered is clinical disorders featuring prominent frontal dysfunction occurring later in life. We identified three recent studies that contribute to this newer line of inquiry by investigating Parkinson's disease, Frontotemporal dementia, and Schizophrenia. As a follow-up to their research examining how decline of fluid intelligence contributes to executive functioning deficits after focal frontal lobe lesions (Roca et al. 2010, as discussed above), Roca, Duncan, and colleagues used identical methodologies to investigate intelligence and executive functioning in individuals with Parkinson's disease (Roca et al. 2012) and Frontotemporal dementia (Roca et al. 2013), two neurodegenerative disorders that prominently include cognitive changes in these domains. Though there were slight differences, overall results from the Parkinson's disease and Frontotemporal dementia studies were highly consistent with their focal frontal lesion data (Roca et al. 2010): some frontal deficits (as measured with the "classic" executive tasks, i.e., Wisconsin Card Sorting Test and verbal fluency) were entirely explained by fluid intelligence loss, while others (multitasking and theory of mind) were not. Interestingly, these researchers propose that the overlap between tests sensitive to frontal impairment and fluid intelligence appears to be on a continuum, and the dissociation between fluid intelligence and frontal deficits depend “on somewhat different frontal regions, with fluid intelligence dependent in particular on lateral and dorsomedial regions (Bishop et al. 2008; Woolgar et al. 2010), whereas more of the anterior frontal cortex is crucial for multitasking and theory of mind” (Roca et al. 2012, p. 2450). These authors’ investigation into executive functioning and intelligence in the presence of neurodegenerative diseases contributes valuable insights and encourages future inquiry into these and related diseases with prominent executive and intelligence involvement. Similar to our discussion of their original study (Roca et al. 2010) above, it is interesting that executive tests most ecologically unique from those of intelligence tests appear to remain significant after controlling for IQ, and collectively these studies highlight the continued difficulty researchers face when selecting assessment tools and attempting to differentiate between two highly overlapping constructs.

In a unique approach, Dichter et al. (2006) examined the relationships between intelligence (WAIS estimated IQ), executive functioning (Trail Making Test, Tower of London, Continuous Performance Task, Wisconsin Card Sorting Test, and a visuospatial working memory task requiring recall of a dot’s location on a screen after a 5-s delay), and P300 event-related potentials (elicited using visual and auditory oddball paradigms) in individuals with Schizophrenia. Estimated IQ was significantly correlated with most executive functioning scores in controls but not in individuals with Schizophrenia. Additionally, for both groups P3 amplitude and latency were significantly correlated with executive functioning, but not with estimated IQ until variance due to executive functioning was removed. Based on their findings, the authors suggest that P3-intelligence relationships may be mediated by executive function abilities and that executive function abilities should be controlled for when examining electrophysiological correlates of intelligence.

Considering all of these studies and their findings, we encourage further examination of changes in executive functioning and intelligence in these and other clinical neurodevelopmental and neurodegenerative disorders, along with continued consideration of the neuroanatomical substrates of age-related cognitive gains and declines (Bugg et al. 2006; Craik and Bialystok 2006; Salthouse 2011).

Additional Considerations

As we have presented here, the primary body of literature examining the relationship between executive functioning and intelligence has consisted of psychometric and/or imaging studies in brain-lesioned patients versus healthy controls, with a closely related secondary body of literature focusing on these relationships in other relevant clinical conditions. Numerous other factors and considerations, however, may hold significant bearing in resolving the ambiguity between executive functioning and intelligence. Although not readily referred to in the current body of literature, we would feel amiss not to acknowledge at least some of these considerations including intellectually gifted children (Arffa 2007), sex differences in higher cognition or frontoparietal network activation (cf. Jung and Haier 2007; Njemanze 2005), and drug treatments that affect executive functioning but not intelligence (cf. Crinella and Yu 1999).

A further consideration pertains to training targeted to improve specific cognitive functions. Three compelling research studies recently examined the relationship between music training, executive functioning, and intelligence (Degé et al. 2011; Moreno et al. 2011; Schellenberg 2011). Some researchers have theorized that executive functions might serve as mediators of intelligence, explaining improvements in IQ associated with learning to play music, and that music training, targeting executive functioning, changes functional brain plasticity and enhances verbal intelligence through high-level cognitive skill transfer. These stances have received both positive (Degé et al. 2011;

Moreno et al. 2011) and negative (Schellenberg 2011) support, though several conceptual and methodological issues remain to be resolved (Hargreaves and Aksentijevic 2011).

Conclusion

So what conclusions can we draw about the relationship between intelligence and executive functioning? The current evidence suggests that as we are still trying to understand the differences between these two complex outcomes, there are some benefits in addressing them as different psychological constructs, despite the fact that their definitions significantly overlap and that they seem to be drawing resources from the same underlying processes. Based on the weaknesses in the literature, we might expect to gain clarity by improving and differentiating the definitions of intelligence and executive functions and their proposed components. Additionally, critical evaluation and thoughtful selection of assessment measures for higher-order functioning in coordination with the application of strong analytic approaches (e.g., latent variables) appear to be key elements to future research in this field. Valuable insights into intelligence versus executive functioning are likely to arise from developments in examining the most evolutionarily advanced (but perhaps least understood) cortical areas and networks, such as the frontoparietal network (e.g., Ramnani and Owen 2004). Furthermore, it seems inevitable that we will ultimately have to adjust our current assumptions about the hierarchical structure of cognitive abilities and psychometric *g* (Haier 2011; Schlinger 2003). For example, considering the amount of variance shared between executive functioning and intelligence (e.g., Barbey et al. 2012b; Schretlen et al. 2000) and their consistent activation of highly similar brain networks, it may not be wise to make a priori assumptions about a hierarchical relationship between the two or that intelligence should always be controlled for before looking at executive functioning (or vice versa).

Along similar lines, it may be more conceptually and practically beneficial to emphasize the differences between intelligence and executive functioning, rather than their similarities. Recently, considerations regarding the differentiating roles of complexity and novelty suggest these are important aspects that may further our understanding of, or differences between, cognitive abilities underlying higher-order cognitive processes (Dumontheil et al. 2011; Duncan et al. 2012; Garlick and Sejnowski 2006; Hampshire et al. 2011), although such considerations have been made more commonly from the perspective of the field of intelligence research, rather than from the perspective of executive functioning research or a balance between the two.

In Fig. 27.2 we attempt to visually represent a purely speculative and theoretically based summary of what we identified to be the current trends in explaining the relationships between the constructs of intelligence and executive functioning. Starting at the bottom of the figure, a three-layered box contains examples of cognitive processes frequently discussed in the conceptualization of executive functioning and intelligence (e.g., problem solving, synthesis, forethought). Some of them were discussed in earlier sections of this chapter (e.g., updating, shifting, problem representation) and several others were not (e.g., monitoring, prospective memory, categorization). Notice that dotted lines mark the separation between the three layers, indicating a certain level of flexibility and mobility between the cognitive processes (or concepts) included in each one of them. Each layer represents a state varying across time (past, current, future). Mental representations and experiences consolidated through knowledge (either educational or experiential, and committed to long-term memory), and skills strengthened via repetition (e.g., ability to synthesize) are represented in the past or "crystallized" states. Examples of mental operations and cognitive processes believed to be actively engaged in responding to the ever-changing environmental demands are included in the layer representing the current or "fluid" states. Future-oriented states involve examples of cognitive processes that might be actively engaged in

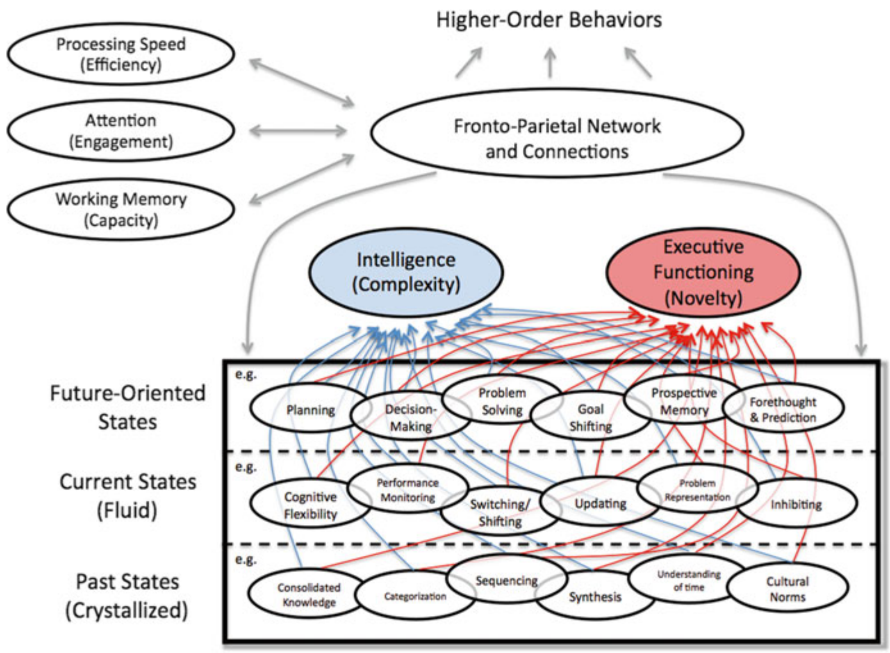


Fig. 27.2 Speculative integration model explaining possible dissociation between executive functioning and intelligence, both as outcomes of common cognitive resources, and the interactions between fluid and crystallized states, as well as states influencing future actions

the mental operations applied to future actions (e.g., prospective memory, decision making). Taken as a group, these processes are continuously interacting with each other at different rates and, depending on the task, are based on goal-directed behavior or environmental demands. The outcomes of these complex and multiple interactions produce singularities, which can be conceptually captured under the labels "Intelligence" and "Executive Functioning." Taking a step forward, we propose to qualify each one of these constructs as being products of differential type of demands, that is, intelligence (as in *g*) fully emerges in situations when the "complexity" of the demands to the system is high, whereas executive functioning emerges in cases in which the demands have a unique "novelty" aspect to them. However, both constructs are outcomes of the same underlying processes, mental operations, and resources.

Furthermore, the processes contained within the time-sensitive states may have a common neural base, the frontoparietal network and its rich cortico-subcortical extensions. Ultimately, this network may serve as a relevant underlying neural support for the generation of behaviors often times referred to be of a "higher order," which can be also categorized as intelligent and/or executive behaviors. Although outside the constraints of this chapter, we cannot neglect three constructs representing basic feed-forward and feedback resources essential to the functioning of the frontoparietal network. These are, *processing speed* (relevant to the efficiency of the system), *attention* (moderating the levels of system engagement), and *working memory* (reduced—solely for the purpose of this theoretically based diagram—to be the resource representing the system's capacity to hold and maintain active the representations needed during all the mental operations outputted via the processing of the brain structures within the frontoparietal network).

Though intelligence, executive functioning, and their corresponding relationships remain elusive, continued research—emphasizing collaborative and integrative approaches—in the areas of lesion studies, psychometrics, neuroimaging, genetics, clinical disorders, neuroplasticity, and numerous other lines of inquiry will most certainly stimulate future advancement in this field.

References

Ackerly, S. (1937). Instinctive, emotional and mental changes following prefrontal lobe extirpation. *American Journal of Psychiatry, 92*, 717–729.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60. doi:10.1037/0033-2909.131.1.30.

Adrover-Roig, D., Sesé, A., Barceló, F., & Palmer, A. (2012). A latent variable approach to executive control in healthy ageing. *Brain and Cognition, 78*(3), 284–299. doi:10.1016/j.bandc.2012.01.005.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*(1), 17–42. doi:10.1007/s11065-006-9002-x.

Anderson, S. W., Damasio, H., Jones, R. D., & Tranel, D. (1991). Wisconsin card sorting test performance as a measure of frontal lobe damage. *Journal of Clinical Experimental Neuropsychology, 13*, 909–922. doi:10.1080/01688639108405107.

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). Cross-sectional and longitudinal patterns of dedifferentiation in late-life cognitive and sensory function: The effects of age, ability, attrition, and occasion of measurement. *Journal of Experimental Psychology: General, 132*(3), 470. doi:10.1037/0096-3445.132.3.470.

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology, 15*(1), 31–36. doi:10.1016/S0887-6177(98)00159-0.

Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology, 22*(8), 969–978. doi:10.1016/j.acn.2007.08.001.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A, 49*(1), 5–28. doi:10.1080/027249896392784.

Baldo, J. V., & Dronkers, N. F. (2006). The role of inferior parietal and inferior frontal cortex in working memory. *Neuropsychology, 20*, 529–538. doi:10.1037/0894-4105.20.5.529.

Balinsky, B. (1941). An analysis of the mental factors of various age groups from nine to sixty. *Genetic Psychology Monographs, 23*, 191–243.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23*, 611–626. doi:10.1037/0012-1649.23.5.611.

Barbey, A. K., Colom, R., & Grafman, J. (2012). Dorsolateral prefrontal contributions to human intelligence. *Neuropsychologia*. (In press) doi:10.1016/j.neuropsychologia.2012.05.017.

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012b). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain, 135*(4), 1154–1164. doi:10.1093/brain/aws021.

Barcelo, F., & Santome-Calleja, A. (2000). A critical review of the specificity of the Wisconsin card-sorting test for the assessment of prefrontal function. *Revista de Neurologia, 30*(9), 855.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive function: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65–94. doi:10.1037/0033-2909.121.1.65.

Bar-On, R., Tranel, D., Denburg, N. L., & Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain, 126*, 1790–1800. doi:10.1093/brain/awg177.

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review, 116*(3), 567. doi:10.1037/a0016262.

Basso, A., De Renzi, E., Faglioni, P., Scotti, G., & Spinnler, H. (1973). Neuropsychological evidence for the existence of cerebral areas critical to the performance of intelligence tasks. *Brain, 96*, 715–728. doi:10.1093/brain/96.4.715.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1), 7–15. doi:10.1016/0010-0277(94)90018-3.

Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *The Journal of Neuroscience, 18*(1), 428–437.

Berlin, L. (2003). *The role of inhibitory control and executive functioning in hyperactivity/ADHD*. Comprehensive summaries of Uppsala dissertations from the Faculty of Social Sciences 120. Uppsala: Acta Universitatis Upsaliensis.

Binet, A., & Simon, T. (1916). The development of intelligence in children (The Binet-Simon scale) translated from articles in *L'Année psychologique from 1905, 1908, 1911* by Elizabeth S. Kite. (Publications of the training school at Vineland New Jersey). Baltimore: Williams & Wilkins Company.

Bishop, S. J., Fossella, J., Croucher, C. J., & Duncan, J. (2008). COMT val158met genotype affects recruitment of neural mechanisms supporting fluid intelligence. *Cerebral Cortex, 18*(9), 2132–2140. doi:10.1093/cercor/bhm240.

Black, F. W. (1976). Cognitive deficits in patients with unilateral war-related frontal lobe lesions. *Journal of Clinical Psychology, 32*, 366–372. doi:10.1002/1097-4679(197604)32:2<366::AID-JCLP2270320234>3.0.CO;2-F.

Blair, C. (2006). Towards a revised theory of general intelligence: Further examination of fluid cognitive abilities as a unique aspect of human cognition. *Behavioral and Brain Sciences, 29*, 145–153. doi:10.1017/S0140525X06419038.

Blair, C. (2010). Fluid cognitive abilities and general intelligence. A life-span neuroscience perspective. In R. M. Lerner (Series Ed.) & W. F. Overton (Volume Ed.), *Handbook of life-span development*: *Vol. 1. Cognition, biology and methods* (pp. 226–258). Hovoken: John Wiley & Sons.

Blair, R. J. R., & Cipolotti, L. (2000). Impaired social response reversal: A case of “acquired sociopathy”. *Brain, 123*, 1122–1141. doi:10.1093/brain/123.6.1122.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26*(2), 571–593. doi:10.1207/s15326942dn2602_3.

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence, 40*(5), 458–469. doi:10.1016/j.intell.2012.05.006.

Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Human Brain Mapping, 25*, 35–45. doi:10.1002/hbm.20128.

Bugg, J. M., Zook, N. A., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain and Cognition, 62*(1), 9–16. doi:10.1016/j.bandc.2006.02.006.

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34*, 263–272. doi: 10.1016/0028-3932(95)00104-2.

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society, 4*(6), 547–558.

Campione, J. C., & Brown, A. L. (1978). Toward a theory of intelligence: Contributions from research with retarded children. *Intelligence, 2*, 279–304. doi:10.1016/0160-2896(78)90020-X.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1–22.

Chiang, M. C., Barysheva, M., Shattuck, D. W., Lee, A. D., Madsen, S. K., Avedissian, C., Klunder, A. D., Toga, A. W., McMahon, K. L., de Zubicaray, G. I., Wright, M. J., Srivastava, A., Balov, N., & Thompson, P. M. (2009). Genetics of brain fiber architecture and intellectual performance. *The Journal of Neuroscience, 29*(7), 2212–2224. doi:10.1523/JNEUROSCI.4184-08.2009.

Collette, F., Hogge, M., Salmon, E., & Van der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience, 139*(1), 209–221. doi:10.1016/j.neuroscience.2005.05.035.

Colom, R., & Thompson, P. M. (2011). Understanding human intelligence by imaging the brain. In T. Chamoroo-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 330–352). Oxford: Wiley-Blackwell.

Colom, R., Haier, R. J., Head, K., Alvarez-Linera, J., Angeles Quiroga, M., Chun Shih, P., & Jung, R. E. (2009). Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence, 37*, 124–135. doi:10.1016/j.intell.2008.07.007.

Colom, R., Karama, S., Jung, R. E., & Haier, R. J. (2010). Human intelligence and brain networks. *Dialogues in Clinical Neuroscience, 12*(4), 489.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences, 10*(3), 131–138. doi:10.1016/j.tics.2006.01.007.

Crinella, F. M., & Yu, J. (1999). Brain mechanisms and intelligence: Psychometric g and executive function. *Intelligence, 27*(4), 299–327. doi:10.1016/S0160-2896(99)00021-5.

Cummings, J. L., & Benson, D. F. (1990). Subcortical mechanisms and human thought. In J. L. Cummings (Ed.), *Subcortical dementia* (pp. 251–260). New York: Oxford University Press.

D’Esposito, M., & Postle, B. R. (1999). The dependence of span and delayed response performance on prefrontal cortex. *Neuropsychologia, 37*, 89–101. doi:10.1016/S0028-3932(99)00021-4.

D’Esposito, M., Cooney, J. W., Gazzaley, A., Gibbs, S. E. B., & Postle, B. R. (2006). Is the prefrontal cortex necessary for delay task performance? Evidence from lesion and fMRI data. *Journal of the International Neuropsychological Society, 12*, 248–260. doi:10.1017/S1355617706060322.

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioral Brain Research, 41*(2), 81–94. doi:10.1016/0166-4328(90)90144-4.

Das, J. P. (1973). Cultural deprivation and cognitive competence. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6, pp. 1–53). New York: Academic.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Needham Heights: Allyn & Bacon.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience, 11*(3), 201–211. doi:10.1038/nrn2793.

Degé, F., Kubicek, C., & Schwarzer, G. (2011). Music lessons and intelligence: A relation mediated by executive functions. *Music Perception, 29*(2), 195–201.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis Kaplan executive function system*. San Antonio: The Psychological Corporation.

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157–173. doi:10.1016/0160-2896(91)90028-C.

Denckla, M. B. (1996). A theory and model of executive function. A neuropsychological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 263–278). Baltimore: P.H. Brookes.

Detterman, D. K. (1986). Human intelligence is a complex system of separate processes. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 57–61). Norwood: Ablex Publishing Corporation.

Detterman, D. K. (2000). General intelligence and the definition of phenotypes. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The nature of intelligence, Novartis foundation symposium 233* (pp. 136–148). Chichester: Wiley.

Diamond, A., Prevor, M. B., Callendar, G., & Druin, D. P. (1997). Prefrontal cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development, 62*(4), 1–208.

Dichter, G. S., van der Stelt, O., Boch, J. L., & Belger, A. (2006). Relations among intelligence, executive function, and P300 event related potentials in schizophrenia. *The Journal of Nervous and Mental Disease, 194*(3), 179–187. doi:10.1097/01.nmd.0000202490.97425.de.

Dumontheil, I., Thompson, R., & Duncan, J. (2011). Assembly and use of new task rules in fronto-parietal cortex. *Journal of Cognitive Neuroscience, 23*(1), 168–182. doi:10.1162/jocn.2010.21439.

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology, 59*, 2–27. doi:10.1080/17470210500260674.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*(10), 475–483. doi:10.1016/S0166-2236(00)01633-7.

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychological, 33*, 261–268. doi:10.1016/0028-3932(94)00124-8.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*, 257–303. doi:10.1006/cogp.1996.0008.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F. N., & Emslie, H. (2000). A neural basis for general intelligence. *Science, 289*, 457–460. doi:10.1126/science.289.5478.457.

Duncan, J., Schramm, M., Thompson, R., & Dumontheil, I. (2012). Task rules, working memory, and fluid intelligence. *Psychonomic Bulletin and Review, 19*(5), 864–870. doi:10.3758/s13423-012-0225-y.

Eslinger, P. L., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology, 35*, 1731–1741. doi:10.1212/WNL.35.12.1731.

Esposito, G., Kirkby, B. S., van Horn, J. D., Ellmore, T. M., & Berman, K. F. (1999). Context-dependent, neural system-specific neurophysiological concomitants of ageing: Mapping PET correlates during cognitive activation. *Brain, 122*, 963–979. doi:10.1093/brain/122.5.963.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., & Glisky, M. L. (1999). Executive functioning in preschool children: Performance on A-not-B and other delayed response format tasks. *Brain and Cognition, 41*(2), 178–199. doi:10.1006/brcg.1999.1117.

Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A., & Moehr, M. J. (2011). Executive control and dimensions of problem behaviors in preschool children. *Journal of Child Psychology and Psychiatry, 52*(1), 33–46. doi:10.1111/j.1469-7610.2010.02265.x.

Estes, W. K. (1986). Where is intelligence? In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 63–67). Norwood: Ablex Publishing Corporation.

Floyd, R. G., Bergeron, R., Hamilton, G., & Parra, G. R. (2010). How do executive functions fit with the Cattell–Horn–Carroll model? Some evidence from a joint factor analysis of the Delis–Kaplan executive function system and the Woodcock–Johnson III tests of cognitive abilities. *Psychology in the Schools, 47*(7), 721–738. doi:10.1002/pits.20500.

Frazer, J. (2012). *The integrative neuropsychological theory of executive-related abilities and component transactions (INTERACT): A novel validation study*. Unpublished doctoral dissertation, University of Victoria, Victoria, Canada.

Freeman, F. S. (1955). *Theory and practice of psychological testing*. New York: Holt, Rinehart, and Winston.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology. General, 133*(1), 101. doi:10.1037/0096-3445.133.1.101.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hweitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*, 172–179.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137*(2), 201. doi:10.1037/0096-3445.137.2.201.

Fuster, J. M. (1988). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe*. New York: Raven.

Fuster, J. M. (2000). Executive frontal functions. *Experimental Brain Research, 133*(1), 66–70. doi:10.1007/s002210000401.

Garcia-Barrera, M. A., Kamphaus, R. W., & Bandalos, D. (2011). Theoretical and statistical derivation of a screener for the behavioral assessment of executive functions in children. *Psychological Assessment, 23*, 64–79. doi:10.1037/a0021097.

Garcia-Barrera, M., Frazer, J., & Areshenkoff, C. (2012). Theoretical derivation and empirical validation of an integrative neuropsychological theory of executive-related abilities and component transactions IINTERACT) [Abstract]. *Journal of the International Neuropsychological Society, 18*(S2), 61. doi:10.1017/S1355617712001063.

Garcia-Barrera, M. A., Karr, J. E., & Kamphaus, R. W. (2013). Longitudinal applications of a behavioral screener of executive functioning: Assessing factorial invariance and exploring latent growth. *Psychological Assessment, 25*(4), 1300–1313. doi:10.1037/a0034046.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Garlick, D., & Sejnowski, T. J. (2006). There is more to fluid intelligence than working memory capacity and executive function. *Behavioral and Brain Sciences, 29*, 134–135. doi:10.1017/S0140525X0630903X.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31. doi:10.1037/0033-2909.134.1.31.

Gignac, G., Vernon, P. A., & Wickett, J. C. (2003). Factors influencing the relationship between brain size and intelligence. In E. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 93–106). Kidlington/Oxford: Elsevier Science, Ltd.

Gläscher, J., Tranel, D., Paul, L. K., Rudrauf, D., Rorden, C., Hornaday, A., Grabowski, T., Damasio, H., & Adolphs, R. (2009). Lesion mapping of cognitive abilities linked to intelligence. *Neuron, 61*(5), 681–691. doi:10.1016/j.neuron.2009.01.026.

Gläscher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., & Adolphs, R. (2010). Distributed neural system for general intelligence revealed by lesion mapping. *Proceedings of the National Academy of Sciences, 107*(10), 4705–4709. doi:10.1073/pnas.0910397107.

Gläscher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., Paul, L. K., & Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences, 109*(36), 14681–14686. doi:10.1073/pnas.1206608109.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., 3rd, Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., & Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 101*(21), 8174–8179. doi:10.1073/pnas.0402680101.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum Editor (Ed.), *Handbook of physiology: The nervous system* (Vol. 5, pp. 373–417). Bethesda: American Physiological Society.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316–321. doi:10.1038/nn1014.

Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., Loughead, J., Connolly, J. J., Qiu, H., Mentch, F. D., Abou-Sleiman, P. M., Hakonarson, H., & Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology, 26*(2), 251. doi:10.1037/a0026712.

Haier, R. J. (2011). Biological basis of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 351–370). New York: Cambridge University Press.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2004). Structural brain variation and general intelligence. *NeuroImage, 23*(1), 425–433. doi:10.1016/j.neuroimage.2004.04.025.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2005). The neuroanatomy of general intelligence: Sex matters. *NeuroImage, 25*(1), 320–327. doi:10.1016/j.neuroimage.2004.11.019.

Hampshire, A., Thompson, R., Duncan, J., & Owen, A. M. (2011). Lateral prefrontal cortex subregions make dissociable contributions during fluid reasoning. *Cerebral Cortex, 21*(1), 1–10. doi:10.1093/cercor/bhq085.

Hargreaves, D. J., & Aksentijevic, A. (2011). Music, IQ, and the executive function. *British Journal of Psychology, 102*(3), 306–308. doi:10.1111/j.2044-8295.2011.02029.x.

Harlow, J. M. (1848). Passage of an iron rod through the head. *Boston Medical and Surgical Journal, 39*, 389–393.

Harlow, J. M. (1868). Recovery from the passage of an iron bar through the head. *Publications of the Massachusetts Medical Society, 2*, 327–347. doi:10.1177/0957154X9300401407.

Hebb, D. O. (1945). Man's frontal lobes: A critical review. *Archives of Neurology and Psychiatry, 54*, 10–24.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hebb, D. O., & Penfield, W. (1940). Human behavior after extensive bilateral removal from the frontal lobes. *Archives of Neurology and Psychiatry, 44*(2), 421.

Horn, J. L., & Cattell, R. B. (1967). Age differences in cognitive capabilities. *Acta Psychologica, 26*, 107–129. doi:10.1093/geronj/21.2.210.

Horner, M. D., Flashman, L. A., Freides, D., Epstein, C. M., & Bakay, R. A. (1996). Temporal lobe epilepsy and performance on the Wisconsin card sorting test. *Journal of Clinical Experimental Neuropsychology, 18*, 310–313. doi:10.1080/01688639608408285.

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology, 35*(1), 20–36. doi:10.1080/87565640903325691.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function:

Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017–2036. doi:10.1016/j.neuropsychologia.2006.01.010.

Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology, 22*(4), 508. doi:10.1037/0894-4105.22.4.508.

Isingrini, M., & Vazou, F. (1997). Relation between fluid intelligence and frontal lobe in older adults. *International Journal of Aging and Human Development, 45*, 99–109. doi:10.2190/WHWX-YNVB-079V-2L74.

Janowski, J. S., Shimamura, A. P., & Squire, L. R. (1989). Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology, 17*, 3–11.

Johnson, W., & Bouchard, T. J., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence, 33*(4), 393–416. doi:10.1016/j.intell.2004.12.002.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences, 30*(2), 135–154. doi:10.1017/S0140525X07001185.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review, 17*(3), 213–233. doi:10.1007/s11065-007-9040-z.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86*, 199–225. doi:10.1016/0001-6918(94)90003-5.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review, 9*(4), 637–671. doi:10.3758/BF03196323.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle. *Psychological Bulletin, 131*, 66–71. doi:10.1037/0033-2909.131.1.66.

Krikorian, R., & Bartok, J. (1998). Developmental data for the Porteus Maze test. *The Clinical Neuropsychologist, 12*, 305–310. doi:10.1076/clin.12.3.305.1984.

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59–80. doi:10.1348/026151003321164627.

Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology, 17*, 281–297. doi:10.1080/00207598208247445.

Lezak, M. D., Howieson, D., Bigler, E., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York: Oxford University Press.

Luders, E., Narr, K. L., Thompson, P. M., & Toga, A. W. (2009). Neuroanatomical correlates of intelligence. *Intelligence, 37*(2), 156–163. doi:10.1016/j.intell.2008.07.002.

Luria, A. R. (1973). *The working brain*. New York: Basic Books.

Martinez, K., Villalon-Reina, J., Kessel, D., Joshi, A., Pineda, J. A., Jahanshad, N., Nir, T., Eschenburg, K., Roman, F. J., Burgaleta, M., Solana, A. B., Angeles Quiroga, M., Colom, R., & Thompson, P. M. (2013). Exploratory factor analysis of brain networks reveals sub-networks related to cognitive performance. *International Symposium on Biomedical Imaging, 10*, 608–611. doi:10.1109/ISBI.2013.6556548.

McAuley, T., & White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology, 108*(3), 453–468. doi:10.1016/j.jecp.2010.08.009.

McDaniel, M. A. (2005). Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence, 33*(4), 337–346. doi:10.1016/j.intell.2004.11.005.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10. doi:10.1016/j.intell.2008.08.004.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in *children* with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240–260. doi:10.1006/jecp.1999.2516.

Miller, E. K. (2000). The neural basis of top-down control of visual attention in the prefrontal cortex. In S. Monsell & J. Driver (Eds.), *Attention and performance* (p. 18). Cambridge, MA: MIT Press.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167–202. doi:10.1146/annurev.neuro.24.1.167.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development, 13*(3), 395–423. doi:10.1080/15248372.2011.585478.

Milner, B. (1963). Effects of different brain lesions on cards sorting: The role of the frontal lobes. *Archives of Neurology, 9*, 90–100. doi:10.1001/archneur.1963.00460070100010.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. doi:10.1006/cogp.1999.0734.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science, 22*(11), 1425–1433. doi:10.1177/0956797611416999.

Muller, N. G., Machado, L., & Knight, R. T. (2002). Contributions of subregions of the prefrontal cortex to working memory: Evidence from brain lesions in humans. *Journal of Cognitive Neuroscience, 14*, 673–686. doi:10.1162/08989290260138582.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77. doi:10.1037/0003-066X.51.2.77.

Neylan, T. C. (1999). Frontal lobe function: Mr. Phineas Gage's famous injury. *The Journal of Neuropsychiatry and Clinical Neurosciences, 11*(2), 280–281.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 241–268. doi:10.3758/s13415-011-0083-5.

Njemanze, P. C. (2005). Cerebral lateralization and general intelligence: Gender differences in a transcranial Doppler study. *Brain and Language, 92*, 234–239. doi:10.1016/j.bandl.2004.06.104.

Osório, A., Cruz, R., Sampaio, A., Garayzábal, E., Martínez-Regueiro, R., Gonçalves, Ó. F., Carracedo, Á., & Fernández-Prieto, M. (2012). How executive functions are related to intelligence in Williams syndrome. *Research in Developmental Disabilities, 33*(4), 1169–1175. doi:10.1016/j.ridd.2012.02.003.

Parkin, A. J., & Java, R. I. (1999). Deterioration of frontal lobe function in normal aging: Influences of fluid intelligence versus perceptual speed. *Neuropsychology, 13*(4), 539. doi:10.1037/0894-4105.13.4.539.

Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. In N. Krasnegor, R. Lyon, & P. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 265–281). Baltimore: Paul H. Brookes.

Pennington, B. F., Bennetto, L., McAleer, O., & Roberts, R. J. (1996). Executive functions and working memory. Theoretical and measurement issues. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 327–348). Baltimore: P.H. Brookes.

Prabhakaran, V., Smith, J. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's progressive matrices test. *Cognitive Psychology, 33*(1), 43–63. doi:10.1006/cogp.1997.0659.

Ptito, A., Crane, J., Leonard, G., Amsel, R., & Caramanos, Z. (1995). Visual-spatial localization by patients with frontal lobe lesions invading or sparing area 46. *Neuroreport, 6*, 45–48.

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience, 5*, 184–194. doi:10.1038/nrn1343.

Raven, J. C. (1938). *Guide to progressive matrices*. London: H. K. Lewis.

Raven, J. C., Court, J. H., & Raven, J. (1988). *Manual for Raven's progressive matrices and vocabulary scales*. London: H. K. Lewis.

Rezai, K., Andreasen, N. C., Alliger, R., Cohen, G., Swayze, V., II, & O'Leary, D. S. (1993). The neuropsychology of the prefrontal cortex. *Archives of Neurology, 50*, 636–642. doi:10.1001/archneur.1993.00540060066020.

Roberts, A. C., Robbins, T. W., & Weiskrantz, L. (1998). *The prefrontal cortex: Executive and cognitive functions*. Oxford: Oxford University Press.

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain, 133*(1), 234–247. doi:10.1093/brain/awp269.

Roca, M., Manes, F., Chade, A., Gleichgerrcht, E., Gershanik, O., Arévalo, G. G., Torralva, T., & Duncan, J. (2012). The relationship between executive functions and fluid intelligence in Parkinson's disease. *Psychological Medicine, 42*(11), 2445. doi:10.1017/S0033291712000451.

Roca, M., Manes, F., Gleichgerrcht, E., Watson, P., Ibáñez, A., Thompson, R., Torralva, T., & Duncan, J. (2013). Intelligence and executive functions in frontotemporal dementia. *Neuropsychologia, 51*(4), 725–730. doi:10.1016/j.neuropsychologia.2013.01.008.

Sadeh, S. S., Burns, M. K., & Sullivan, A. L. (2012). Examining an executive function rating scale as a predictor of achievement in children at risk for behavior problems. *School Psychology Quarterly, 27*(4), 236. doi:10.1037/spq0000012.

Salthouse, T. A. (2001). Structural models of the relations between age and measures of cognitive functioning. *Intelligence, 29*, 93–115. doi:10.1016/S0160-2896(00)00040-4.

Salthouse, T. A. (2011). Neuroanatomical substrates of age-related cognitive decline. *Psychological Bulletin, 137*(5), 753. doi:10.1037/a0023262.

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review, 15*(2), 364–371. doi:10.3758/PBR.15.2.364.

Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia, 29*(12), 1241–1249. doi:10.1016/0028-3932(91)90037-9.

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology, 102*(3), 283–302. doi:10.1111/j.2044-8295.2010.02000.x.

Schlinger, H. D. (2003). The myth of intelligence. *The Psychological Record, 53*(1), 2.

Schretlen, D., Pearlson, G. D., Anthony, J. C., Aylward, E. H., Augustine, A. M., Davis, A., & Barta, P. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society, 6*(1), 52–61. doi:10.1017/S1355617700611062.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*, 727–741. doi:10.1093/brain/114.2.727.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist, 42*(2), 137. doi: 10.1037/0003-066X.42.2.137.

Spearman, C. (1904). “General intelligence,” objectively determined and measured. *The American Journal of Psychology, 15*(2), 201–292.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*, 114–127. doi:10.1037/0022-0663.89.1.114.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (2000). The concept of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 3–15). New York: Cambridge University Press.

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood: Ablex Publishing Corporation.

Stoddard, G. D. (1943). *The meaning of intelligence*. New York: Macmillan.

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin, 95*(1), 3. doi:10.1037/0033-2909.95.1.3.

Stuss, D. T., Benson, D. F., Kaplan, E. F., Weir, W. S., Naeser, M. A., Lieberman, I., & Ferrill, D. (1983). The involvement of orbitofrontal cerebrum in cognitive tasks. *Neuropsychologia, 21*, 235–248. doi:10.1016/0028-3932(83)90040-4.

Swanson, H. L. (1999). Reading comprehension and working memory in learning-disabled readers: Is the phonological loop more important than the executive system? *Journal of Experimental Child Psychology, 72*, 1–31. doi:10.1006/jecp.1998.2477.

Terman, L. M. (1921). A symposium: Intelligence and its measurement. *Journal of Educational Psychology, 12*, 127–133. doi:10.1037/h0064940.

Teuber, H. L. (1959). Some alterations in behavior after cerebral lesion in man. In A. D. Bass (Ed.), *Evolution of nervous control from primitive organisms to man* (pp. 157–194). Washington, DC: America Association for the Advancement of Science.

Teuber, H. L. (1964). The riddle of frontal lobe function in man. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behavior* (pp. 410–444). New York: McGraw-Hill.

Thomson, G. H. (1951). *The factorial analysis of human ability* (5th ed.). London: University of London Press.

Toga, A. W., & Thompson, P. M. (2005). Genetics of brain structure and intelligence. *Annual Review of Neuroscience, 28*, 1–23. doi:10.1146/annurev.neuro.28.061604.135655.

Tranel, D., Manzel, K., & Anderson, S. W. (2008). Is the prefrontal cortex important for fluid intelligence?: A neuropsychological study using matrix reasoning. *The Clinical Neuropsychologist, 22*(2), 242–261. doi:10.1080/13854040701218410.

Tsuchida, A., & Fellows, L. K. (2009). Lesion evidence that two distinct regions within prefrontal cortex are critical for n-back performance in humans. *Journal of Cognitive Neuroscience, 12*, 2263–2275. doi:10.1162/jocn.2008.21172.

Tsujimoto, S. (2008). The prefrontal cortex: Functional neural development during early childhood. *The Neuroscientist, 14*(4), 345–358. doi:10.1177/1073858408316002.

van der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*(4), 842. doi:10.1037/0033-295X.113.4.842.

Volle, E., Kinkingnéhun, S., Pochon, J., Mondon, K., de Schotten, M. T., Seassau, M., Duffau, H., Samson, Y., Dubois, B., & Levy, R. (2008). The functional architecture of the left posterior and lateral prefrontal cortex in humans. *Cerebral Cortex, 18*, 1093–1103. doi:10.1093/cercor/bhn010.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience, 3*(4), 255–274. doi:10.3758/CABN.3.4.255.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.

Weinstein, S., & Teuber, H. L. (1957). Effects of penetrating brain injury on intelligence test scores. *Science, 125*, 249–263. doi:10.1126/science.125.3256.1036.

Wickett, J. C., Vernon, P. A., & Lee, D. H. (2000). Relationships between factors of intelligence and brain volume. *Personality and Individual Differences, 29*(6), 1095–1122. doi:10.1016/S0191-8869(99)00258-5.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*(3), 436–452. doi:10.1016/j.jecp.2010.08.008.

Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex-basal ganglia system in primates. *Critical Reviews in Neurobiology, 10*, 317–356. doi:10.1615/CritRevNeurobiol.v10.i3-4.30.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Itasca: Riverside Publishing.

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., Roca, M., Antoun, N., Manes, F., & Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences, 107*(33), 14899–14902. doi:10.1073/pnas.1007928107.

The Evolution of Intelligence: Implications for Educational Programming and Policy

28

Christopher Jones, Peggy L. Tarpley, and Douglas Blancero

When the opportunity arose to write this chapter, we felt it would be a great opportunity to write about why teachers often reject changing based on research-guided knowledge, in this case specifically about intelligence but broadly about education. As a team of academics, practitioners, and coaches, we have worked in schools for over 25 years. The focus on the rejection of research as primary to what we have observed in schools over time was an easy one. The issue is complex with tendrils that reach out and affect all aspects of school reform.

Over the years we have identified several factors affecting the growth and improvement of schools. This chapter addresses several of them. A common thread running through all of these factors is change. How we, the academics we, the academics and trainers/coaches, and the schools, the administrators, teachers, students, and families perceive and accept change affects how both a school and the members of the school's academic community function, grow, succeed, or fail.

C. Jones
Longwood University, Hull Building,
201 High Street, Farmville, VA 23909, USA

P.L. Tarpley
Department of Education and Special Education,
Longwood University, 201 High Street,
Farmville, VA 23909, USA

D. Blancero (✉)
JP Associates, 7 Amherst Street,
Red Hook, NY 12571, USA
e-mail: dblancero@jponline.com

Carnine Revisited

Educators many times seem reluctant to embrace research diametrically opposed to their opinions or previous learning. Understanding the complex relationship between teaching and learning is often convoluted. Many teachers work incredibly hard to educate children yet produce less than stellar results. In education, we can often find excuses why children do not learn that are not aligned with the instruction. Avoiding culpability insulates us from internal and external criticism and allows us to continue to use ineffective methods. While we were having extensive conversations on this very topic, we wondered what it would take to get teaching to become more of a profession that valued research. After a lot of discussion on how to approach this chapter, we decided to revisit Douglas Carnine's reflections on our profession.

During the winter of 2000, Douglas Carnine wrote an article titled, *Why education experts resist effective practices (And what it would take to make education more like medicine)*. In his article he lamented about the teaching profession. He questioned why teachers embrace methods that do not work. He wondered why the findings of the national reading panel were still

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_28,

excluded from practice in many K-3 classrooms. He was also critical of college professors who are charged with the task of producing effective teachers. He included Diane Ravitch's trip to the intensive care unit and her visions of what her healthcare would be like if the physicians used the same methodology that education professors used. Carnine called on our profession to embrace research at an individual teacher level. It seems like it is a good time to revisit Carnine's article to see how far we have come. It has been 15 years since Carnine's article, and we have seen many reforms come and go with mixed results (e.g., Reading First; Reading First Implementation Evaluation: Final Report 2008). We have watched our science, reading, and math scores stay flat, while other nations thrive (http://www.oecd.org/pisa/). We now see new initiatives, Race To the Top, aligned with new standards (Common Core) with similar goals. But the question continues to trouble us, why are we seeing academic success in other countries, yet we continue to struggle with literacy and basic math skills in the United States? Has the United States reached their plateau with regard to academic performance or is it something else? Is our lack of student achievement aligned more with the student variables or teacher variables? If our academic stagnation is caused by student variables, then indeed we might be doing as well as we can.

These are tough questions that can possibly be answered by historians, preachers, economists, and a number of other groups. We take the position that instead of placing the primary blame on outside variables, maybe we should first examine the disposition of our teachers and their implementation of best practices.

It is our opinion that we have not reached our pinnacle. It is our belief that our stagnant academic performance is in part a by-product of the rejection of evidence-based practices. It may be tempting to wag one's disapproving finger at teachers and yell "Shame on you teachers," but it might be more productive to open a dialogue focusing on possible reasons why there is a rejection of the evidence in our field. If we do not identify the causation of our rejection of evidence and propose possible treatments to modify this characteristic, we are afraid the previous chapters in this text might not have an impact on student learning.

Casting blame, devoid of any reflection, on how we got to this position will not result in an increase in student academic performance or an increase in educators embracing evidence. It is more important for educators to reflect on how we got here rather than who is at fault. Reflecting on the original purpose of education is as good a place as any to start. Thomas Jefferson proclaimed:

> Now let us see what the present primary schools cost us, on the supposition that all the children of 10, 11 and 12 years old are, as they ought to be, at school: and, if they are not, so much the work is the system; for they will be untaught, and their ignorance & vices will, in future life cost us much dearer in their consequences, than it would have done, in their correction, by a good education (Ford 1892–99).

School is an intervention for ignorance. If our primary intervention for ignorance is ineffective, then we have to modify it.

Shut Down the Schools?

It has been 50 years since the verdict was passed down on Brown versus Board of education. But the ramifications for school systems rejecting the court's finding are still felt by many students who, through no fault of their own, experienced very little or no education for over 7 years due to the closing of certain public schools. During this volatile time, many students in Prince Edward County, Virginia, received their education in church basements, private homes, or other makeshift facilities, while some just dropped out of school. While it is not the focus of this chapter to revisit those dark times, it is important to determine what was lost by those students who dropped out. Conventional wisdom would say that they lost their opportunity; they lost their chance to be competitive. Would our history be very different if our educators focused on the *evidence* of effective teaching and Jefferson's warnings about "future life cost" rather than politics and fear?

It is true that many believed, and still do, that it is a child's history, upbringing, or genetics that is the indicator of their future academic success. Many educators, at least intellectually, reject the notion of academic predestination. If education is indeed the great equalizer (Mann 1848), then it is safe to say that many of the students' lives were made much more difficult because of the politics that resulted in the closing of schools. Aside from politics and background, we must have evidence that highlights the importance of effective instruction.

Hattie (2009) has shown that contributions from the teacher, school, and curricula have a significant impact on student learning. Clearly, how a child gets educated has a greater impact than the variables they bring to school. Lemov (2010) identified 49 techniques that improve student achievement that teachers use. We know that if we have excellent delivery of phonics instruction, we will have a greater impact on literacy skill development (Hattie 2009; National Reading Panel 2000). If we combine systematic explicit phonics instruction ($d=0.6$) with delivery techniques that support student achievement (Lemov 2010), then we can positively impact student literacy achievement. And while this is all very encouraging, remember Chall reported the same findings in 1967. One has to wonder that if we were a profession focused more on Jefferson's vision and the literacy achievement evidence, would we have locked the doors in Prince Edward County?

The Importance of School

It is safe to say that as long as we continue to value opinions, politics, and rhetoric over evidence, the findings of Coleman and Jencks will continue to influence our teachers. Coleman et al. (1966) asserted that student's academic success is based on their background and social standing. While Jencks et al. (1972) mirrored many of Coleman's findings by stating that schools did very little to lessen the achievement gap between the haves and the have-nots, Jencks also reported that student achievement was a by-product of family background and that school reform had little or no influence on student achievement. While subsequent research has validated the importance of an effective teacher using evidence-based procedures (Hattie 2009; Marzano 2003), the importance of education will still be called into question until we quantify and disseminate the value of an efficient and effective teacher using best practices.

Fortunately, a great deal of research shows that schools and more specifically teachers have a significant influence on student achievement. Nye et al. (2004) reported that students who receive instruction from highly effective teachers score between 13 and 18 percentile points higher in reading and math. Marzano (2003) found that what goes on in the classroom accounted for 20 % of the variance in student achievement which was twice as much as the Coleman study (Marzano 2003). Hattie's (2009) synthesis of over 800 meta-analyses relating to achievement reported the variables that produce many of the greatest effect sizes that were teacher-related variables (e.g., providing formative evaluation: ES 0.90, acceleration: ES 0.88, microteaching: ES 0.88, teacher will clarity: ES 0.75, teachers challenging students: ES 0.66, high expectations: ES 0.53). Clearly, teachers using best practices can invoke great change in student performance.

School reform should be a relatively simple task. All we have to do is identify those variables that produce the greatest gains, replicate them in classrooms across the United States, and we should see dramatic increases in student performance. It just does not seem that difficult. Within a matter of months, we should be able to introduce into the classrooms those variables that have the greatest chance for student improvement and provide teachers with specialized professional development that builds their competencies in best practices. We can then provide evidence of our success through ongoing formative progress monitoring.

Clearly, this has not been the case, so where did the process break down? If we know that certain variables produce greater student achievement, then those variables should be taught in higher education teacher preparation programs and during the in-school professional development training sessions. Unfortunately, this line of logic is often broken. There is a disconnect sometimes between the needs of the school divisions/districts, the professional development offered to the teachers (we are not yet even considering the additional aspect of "how" we assess, analyze, and identify what professional development is needed or how it should be delivered in the school setting—stand up training, coaching, etc.), and what is taught in teacher preparation. In order to rectify this, we have seen lucrative grants such as Reading First attempt to impose Scientifically Based Reading Research (SBRR) paired with reading "Coaches." These types of grants produced, at best, mixed results (Reading First Implementation Evaluation: Final Report 2008). On paper Reading First seemed to have a real chance to make significant progress on the war against illiteracy. The grants required the use of SBRR and brought reading coaches into the building to assist teachers with their implementation of the SBRR. While there are many variables that caused problems with the final results of Reading First, one was the top-down approach to having teachers implement SBRR. Many of the methods included in the SBRR stood in stark contrast with the preservice training teachers received. Seasoned teachers also rejected the SBRR on grounds of autonomy or opinion. It became apparent that each individual teacher must understand the nature of evidence-based practices and their influences on student achievement. If this condition is not met, then the intervention will not be successful. It became obvious that in order for best practices to be embraced by many teachers, the teachers needed to value research.

Teachers Who Value Research

Whole Language Versus Phonics

Nothing evokes more heated dialogue in education than the reading wars. Chall and Adams advocated for phonics approaches, while Chomsky and Goodman advocated for more whole language methods Chall (1996). The rhetoric behind these two approaches was passionate and colorful but often failed to cite the empirical evidence that supports their use. Often the selection of an approach came down to what was taught to that individual teacher at the university and college level. These debates went on for decades and succeeded in aligning teachers into different camps. Professors from higher education programs that prepared teachers often found themselves in a position of having to choose which approach they were going to use in their programs. The choice often was based on the professor's preferences, previous institution instructional practices, or current trends. It was not until the National Reading Panel identified the five major components of good solid reading instruction (phonemic awareness, alphabetic principle, fluency with text, vocabulary, comprehension) that we began to embrace the need for good solid systematic explicit phonics instruction. Or did we? Even in the face of the massive NRP finding, we still saw a proliferation of constructivists' literacy programs. It seems that when we get close to accepting an approach as having evidence to support its use, we repackage old unproven methods and present them as if they were something new. The repackaging of unproven methods fulfills the need to try something new but rarely produces desired results. It appears that simply mandating best practices is not going to work. Teachers must enter the field with an understanding and skill set that allows them to interpret what constitutes best practices.

What Is Done in Teacher Preparation Matters

Great atrocities were often committed under noble pretenses. When educators make curricular decisions based on affect, opinion, or tradition, they are bound to make catastrophic mistakes. And even though many individuals do not feel like they are committing atrocities, generation after generation of at-risk students who do not possess basic reading skills or basic mathematics skills is a testament to our failure. Sending teachers into the field who are not well versed in best practices but who possess a romantic idea of what constitutes good solid instruction is going to perpetuate continued student academic failures. The simple truth is that many of our struggling children will not benefit from the methods aligned with romantic ideas of what our teacher candidates consider good solid instruction. Good intentions aside, the instructional choices that we are making contribute to the continued underperformance of at-risk students. We have a tendency to blame variables beyond our control for the student's academic failures. Excuses such as their home environment or social economic status, clearly these have an influence on their achievement, but these are variables beyond our control. The professional teacher would be better off focusing on variables within their control or influence. The most important variable a teacher has control over is their instruction. But there is an appeal for attributing student failure to variables beyond the teacher's control or influence, and that is personal accountability. Attributing student academic failure to variables beyond the individual teacher's control often eliminates the need for instructional adaptation or change. One of the problems with aligning student failure with variables beyond the teacher's control is it legitimizes student failure. Even if the outside influences were correctly identified as reasons for academic failure, causation is not an excuse for continued academic failure. If a child comes to us with an oral language deficit caused by their lack of verbal interactions, it is better for us to seek evidence-based methods that assist the teacher in building the student's oral language vocabulary than to use their background as an excuse for academic failure. Understanding the cause of the student's oral language deficit does not let us off the hook. We still need to address the child's oral language deficits and phonological awareness deficits prior to beginning their reading instruction. The teacher who is well versed in research will know the methods that will effectively address the student's academic weakness. The teacher who is not well versed, or is more aligned with the art of teaching rather than the science of teaching, could continue to provide instruction that does not meet the student's academic needs. The sad truth is we often deny effective interventions either through our own ignorance, disposition, or just plain arrogance.

As a profession, when we reject evidence-based practices, we fail those children who are initially excited, optimistic, and want to learn. And while children come to us with huge academic deficits, with the proper instruction, they can catch up, but if that child is taught by an educator who does not value research, it is a catastrophe waiting to happen.

Art of Teaching Versus Science of Teaching

Carnine (2000) wrote that teacher preparation programs have a tremendous influence on teacher development. They are incredibly influential in the professional organizations that invoke academic policies. There was a time when teacher education programs were justified in their independent approaches. Our teachers taught students who excelled on measures of science, reading, and math. Our student's performance on these measures justified what we were doing in the schools and in turn what we were doing in higher education. And while we continue to produce some of the world's great minds, recent reports indicate that the United States' academic performance in reading, science, and mathematics is flat compared to students in Shanghai, Singapore, and many other Asian providences (PISA—Organization for economic co-operation and development http://www.oecd.org/pisa/).

It is safe to say that a country's future rests on their ability to produce students who can compete in a global economy. If this conventional wisdom is true, the United States is at great risk for losing their competitive edge. Some individuals would question the importance or even the validity of these assessments. Many times these arguments only confuse teachers. And while arguments for and against assessment spark great debate, we cannot deny the staggering number of illiterate and undereducated students in the United States. The drain on our economy and on our way of life is undeniably apparent. The US Department of Education's National Institute of Literacy in 2013 found 14 % of the adults in the United States could not read; that is over 32 million people. The number jumps to 21 % of the population who reads below a fifth grade level. All one has to do to see the damaging effects of literacy is to enter any prison and view a 65 % illiteracy rate. The true challenge is that in spite of massive reform movements, the National Assessment of Adult Literacy has found no significant change in literacy scores for over a decade (http://nces.ed.gov/naal/kf_demographics.asp).

How do we justify the stagnant growth when we have a wealth of research that validates best practices. One possible explanation is that we reject that research in favor of the art of teaching. The art of teaching often becomes more important than the science of teaching. Stanovich and Stanovich (2003) argued that teachers become much more effective if their skill level in identifying evidence-based practices is fully developed. Pearson (1999) agreed with Carnine that education needs to be more in line with the medical model. It is our professional obligation to utilize best practices that originate from valid research. If teachers want to support teacher autonomy, it is imperative that they earn this privilege by consistently aligning their pedagogy based on evolving research. It is this exact method that elevated the public's respect for individuals in the field of medicine. Unfortunately, many in the field of education seem to reject empirical research preferring evidence from personal observations. While clearly anecdotal reports have their place in assessment, the overreliance on qualitative data often results in a rejection of best practices (Pearson 1999). For example, Beck (1998) observed that many proponents of whole language when faced with the findings of the national reading panel continue to reject systematic phonics instruction in favor of whole language. They did this with a clear conscience because they observed many students coming into their classroom who did not need explicit instruction in the alphabetic principle. They noticed that the students who were from print-rich environment settings had highly developed literacy skills. They believed that all one had to do was to replicate the print-rich environment, and students would become literate. It is not their motivation and effort that comes into question, it is their rejection of the research of best practices that hampers student progress.

What are we willing to do to remain competitive with other countries and address the needs of our most at-risk students? Our understanding of teaching sometimes convinces us that if we can make this content interesting and lively, then it will be effective. And while entertainment is definitely an element of teaching, it is not the whole of teaching. The perception that a good teacher knows intuitively if their students are learning conflicts with the need for accurate progress monitoring.

We want instruction to be lively and engaging, and yes a teacher can gauge the general performance of their students in some cases without assessments. But if these techniques were enough, why do we continue to have students lagging behind their peers. If teachers instinctively know how their students are performing and what constitutes effective instruction, then what reason do we have for student failure? Are teachers turning a blind eye to those students who are underperforming in their classroom? We have to believe that this is not the case. The more likely explanation is that teachers simply do not know what to do when traditional methods are ineffective. As a result, they continue to use ineffective methods hoping for a different result. And while it is easy to blame the teachers for their skill deficits, higher education programs must own their accountability for teacher deficits in understanding what constitutes effective instruction.

Professors Unchecked

Within Checks

Professors of higher education often emerge from their Ph.D. program with a belief that they are the unchecked experts in their classrooms. They possess a body of knowledge that has been influenced by other professors, their experience, and their course work. While many Ph.D. programs in the field of teacher education lay a solid pedagogical and theoretical knowledge base, it is imperative that teacher preparation program professors consistently validate and update their instruction based on evidence. Teacher education professors are generally hard workers, but it is the lack of opportunities for internal and external feedback that could prevent them from relying more on evidence-based methods. If education professors fail to overtly seek feedback, then they will be limited by their shortcomings. All one has to do is to look at the evaluation procedures established by respected professions, and one sees the need for sustained purposeful feedback. The field of medicine provides us with an example where multiple parties provide the checks and balances needed to maintain a high level of effectiveness.

External Checks

Doctors are exposed to various checks and balances embedded within their profession. A doctor's work can be evaluated by internal review boards, patients, and even lawyers. Those checks and balances, often absent in higher education, push those in the medical field to be accountable for their decision. Many times education professors receive very little feedback from their colleagues and even less from students. Unfortunately, if a college professor in teacher preparation programs is receiving very little feedback from their colleagues, then it is also likely that they are receiving very little feedback as to the effectiveness of their instruction from external parties. It is very rare to see a college professor actively seeking feedback on their classroom instruction from individuals who work directly with teachers once they are in the classroom. This lack of communication with professional development providers and school systems prevents professors from receiving very valuable information as to the quality of their instruction and program.

Professional development providers have a unique opportunity to not only become well versed in evidence-based procedures, but they can also see the impact of their professional development session on individual classroom teachers. Many of these organizations have a firm understanding of what constitutes good solid instruction. In higher education we often do not forge these very lucrative partnerships between teacher preparation programs and professional development service providers. Many of these improvement specialists are highly experienced and can work with a wide range of schools. By providing on-site professional development, they can have a direct impact on student learning and often collect much of the evidence that supports best practices (JP Associates website http:www.jponline.com). It would seem that a partnership between school divisions, the on-site professional development providers, and teacher preparation programs would benefit the teacher education candidate. We might wonder what it would be like if our teacher preparation programs were more like the medical residency models. Within the residency model, not only do the candidates benefit from professors' instruction, but they also receive feedback from nurses, patient advocates, family members, and a host of other interested parties. And while the sheer volume of feedback could be overwhelming, it pushes the interns to know their craft. We can only imagine the positive impact on teacher education if school divisions, teacher preparation programs, and on-site professional development organizations all work together in educating the future teacher.

Feedback

Actively seeking feedback is an indication that one wants to perfect their abilities to teach. Reciprocal feedback, where we not only give feedback on projects, tests, responses, etc., but seek feedback on the effectiveness of our methods is critical. Having the confidence to actively seek feedback brings you that much closer to valuing evidence.

A true professional, who is open to feedback, is seeking the truth. If they are interested in how effective their instruction is, then they will be equally interested in what works. Wanting to know what works and being open to others' empirically based ideas will push individuals to embrace evidence-based instruction.

If teacher education professors embrace the concepts associated with evidence-based instruction, they will model these concepts to the teacher candidate. Hattie (2009) reported an effect size of 0.77 for feedback. Teacher to student feedback is important, but equally as important is feedback from student to teacher. In fact, some of the greatest gains in student achievement occur when students provide feedback to their teachers (Hattie 2009). And while teachers report that they actively solicit feedback from their students, that is not always the case.

When asked if teachers accept feedback from students, 70 % of the teachers say they accept it, whereas only 45 % of the students say the teachers accept their feedback (Carless 2006). Students' perception of the acceptance of their feedback is critical. Students are recipients of that instruction, and their feedback is a good indicator of how effective the instruction is. When professors actively seek student feedback, they can obtain a fairly accurate gauge as to the effectiveness of their instruction. More importantly they instill a belief in the teacher candidates that feedback is valuable. We believe teachers who value feedback are more open to research.

When we venture into areas that we are not sure of, areas that are beyond our field of expertise, we are much more likely to become defensive if our assertions are challenged. But it is exactly those times we should be actively seeking feedback. Unfortunately, it is very tempting to pretend and hope that individuals will not discover our weaknesses. What we fail to realize is that by modeling our search for information to clarify meaning, we instill in our candidates that they should actively seek feedback in times of uncertainty. We should work with teachers to actively investigate unfamiliar concepts. Sometimes our narrow vision of professionalism prevents us from seeking feedback. If seeking feedback is one of the first steps in embracing evidence, what would it take to get professionals to actively seek feedback?

Change

At the beginning of this chapter, we briefly addressed change and spoke of the need for an acceptance of research. This might be a good time to take a closer look at change and its role in organizational growth, educator acceptance of research, and higher education changing its approach to teacher prep. Lawrence Lezotte tells us, "Any model of school improvement that is going to be useful to schools must focus explicitly on results, evidence of student learning, and student achievement" (Assembly Required: A Continuous School Improvement System, p. 5, Lezotte). And later on in the same chapter, he writes:

> …the system in place is ideally suited to producing the results the school is currently getting … any change in the desired results, from the current system in place is going to require a change in mission, core beliefs, and core values that underpin the system, especially if the goal is to permanently sustain the desired change. (p. #)

There are at the least two main aspects of change—accepting the need for change and managing the change. When we ask people, in this case teachers, to change what they are doing, even when it is based on data, it is often perceived as a challenge to their personal belief, the hard work they have invested. Sometimes it is the actual act of changing that can be the challenge as much as what the change is.

At one point in our work with schools, our approach was to work with teachers, individually and separate from leadership. We felt if we changed what was happening in the classroom, we could change the school. It soon became evident that unless the people of that school connected and shared, unless the leader was providing a clear vision and supporting that vision effectively, then change was slow, or in the worse cases, change was resisted and repelled.

Changing the culture of the building is essential to shifting mind-sets from the current system in place to one that will produce increased success. Michael Fullan calls this reculturing. This reculturing needs to occur on all levels (higher education, schools, and classrooms). It calls for both a strong foundation in effective instruction (looking at the data) and a transformational approach (relationships and sharing knowledge).

Conclusions

School reform should be a relatively simple task, but recent history indicates otherwise. Throughout this chapter we have identified certain traits that prevent teachers from embracing evidence with the result of limiting our understanding of intelligence and intellectual processes thereby flattening math, science, and reading scores for students. Educators are often reluctant to embrace research that does not support their personal approach to instruction. Stagnant academic performance is a by-product of the rejection of evidenced-based practices. This resistance to accept sound research takes on even more importance when research indicates that contributions from the teacher, school, and curricula have a significant impact on student learning.

It is apparent that it is not one individual at fault; rather it is a collective failure of education. It is due, in part, to our tendency to blame variables beyond our control for the student's academic failures. The variable that is within our control is teacher preparation, ongoing professional development, and effective leadership that manages and leads the change necessary for success.

We should instill in the preservice teacher the notion that we must work as a team to address student achievement issues. Professors should be part of a collaborative group that includes school divisions, professional development providers, and teacher preparation programs. Preservice teachers need to actively seek feedback and embrace evidence-based practices. Teachers need to actively seek evidence that supports the methods that they are using. They should be collecting formative data that supports the continued use of these methods. If the data shows that students are not responding to the intervention, then, the teacher should modify that intervention. For teachers to possess these skills, teacher preparation professors must model and validate the use of evidence in their instruction.

To summarize, teacher rejection of research-based ideas and practices impacting educational decisions for students comes about because of several constructs. One, teachers may find that the research clashes with their opinion or anecdotal observations. Further, this clash may come with the teacher's past learning or teacher preparation program. Two, resistance may come because teachers and administrators find the discomfort of changing from what has always been done requires significant effort and friction. Finally, educators find solace in the excuses that lack of educational progress is somehow the fault of past teachers, the family, or the child himself or herself. Without these excuses, teachers and administrators are smack in the midst of accountability or culpability.

Carnine (2000) writes about the pressure on education from all sides. He suggests that all professions have been at this stage of growth toward a mature profession. The recipients of the less than adequate treatment (whatever profession it may be) and their advocates call for the incorporation of more scientific methodology in the profession. He states further, "A mature profession, by contrast, is characterized by a shift from judgments of individual experts to judgments constrained by quantified data that can be inspected by a broad audience…" (p. 12). While all, pre- and in-service teachers, teacher education faculty, school administrators, and professional development providers, contribute to the resistance to education based on research, they also hold promise to the acceptance and implementation of research-based practices.

Model for Success

We need a partnership between school divisions, on-site professional development providers, and teacher preparation programs based on the medical residency model where teacher professional development is a natural extension of teacher preparation. We learn best by doing. On-site professional development with both leadership and teachers as they are working through their normal day can provide needed constructive feedback and guidance options that might not be seen by those in the "thick of the forest." It can be the training and proving ground for effective practices. Here in the field, each individual teacher can participate in evidence-based practices and observe the influences on student achievement. In this way teachers, coaches/professional development providers, and administration are all contributing to the evidence of best practices.

A change in how schools work needs to be explored. A shift from simply managing schools to leading schools is essential. There needs to be a move from top-down management to two-way communication that allows both administrators and teachers together to explore what works for students while at the same time recognizing decisions need to be made and those decisions will be questioned. We believe that teacher preparation professors should be in the discussion as well. If there is a "fair" process in place for such discussion, consensus, collaboration, and cooperation can be reached. It calls for both a strong foundation in effective instruction (looking at the data) and a transformational approach (relationships and sharing knowledge). This speaks to the reculturing Michael Fullan recommends. Such a reculturing would allow educators to actively investigate unfamiliar concepts, expand their vision of professionalism, and actively seek feedback. It will require changing our mind-set and providing a real vision for education that embraces change and views research as a potential new tool for making decisions based on our students' needs and not ours.

Finally, all of these threads are connected by the concept of change. Avoidance of change, whatever the reason, prevents growth. Resistance to change is one of the primary reasons principals, teachers, students, and practitioners fail at school reform. If schools are going to improve and if we as a country are going to reclaim our position as the premier nation for education, we must embrace change—that means embracing new knowledge and research and effectively applying that information to educate our youth and ourselves along with them.

References

Beck, I. (1998). Understanding beginning reading: A journey through teaching and research. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 11–31). New York: Guilford Press.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219–233.

Carnine, D. (2000). Why education experts resist effective practices (And what it would take to make education more like medicine). *Effective School Practices, 18*(3), 6–14.

Chall, J. S. (1996). *Learning to read: The great debate*. Fort Worth: Harcourt Brace College Publishers.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfield, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Ford, P. L. (Ed.) (1892–99). *The writings of Thomas Jefferson* (10 vols). New York: G.P. Putnam's Sons, 1892–99. http://catalog.hathitrust.org/api/volumes/oclc/2428785.html

Fullan, M. (2001). *Leading in a culture of change*. California: Jossey-Bass.

Hattie, J. A. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. New York, Ny: Routledge.

Jencks, C., Smith, M. S., Ackland, H., Bane, M. J., Cohen, D., Grintlis, H., Heynes, B., & Michelson, S. (1972). *Inequality: A measurement of the effects of family and schooling in America*. New York: Basic Books.

Lemov, D. (2010). Teach like a champion: 49 techniques that put students on the path to college. San Francisco, CA: Jossey Bass.

Lezotte, L. W., & McKee, K. M. (2002). *Assembly required: A continuous school improvement system*. Okemos: Effective Schools Products, Ltd, Doug.

Mann, H. (1848). Annual Report to the Massachusetts Board of Education.

Marzano, R. J. (2003). What works in schools: Translating research into action. Alexandria, Va: ASCD.

National Reading Panel. (2000). http://www.nichd.nih.gov/research/supported/Pages/nrp.aspx/. Accessed 9/24/2014.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large our teacher affects? *Education Evaluation and Policy Analysis, 26*(3), 237–257.

Pearson, P. D. (1999). A historically based review of preventing reading difficulties in young children. Reading Research Quarterly, 34, 231–246.

PISA – Organization for Economic Co-operation and Development. Reading first implementation evaluation: Final report, 2008. http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm.

Stanovich, P. J., & Stanovich, K. E. (2003). Using research and reason in education: How teachers can use scientifically based research to make curricular & instructional decisions. University of Toronto. https://www.nichd.nih.gov/publications/pubs/Pages/using_research_stanovich.aspx.

U.S. Department of Education. (2013, Apr 28). National Institute of Literacy Research.http://www.statisticbrain.com/number-of-american-adults-who-cant-read/.

Part VI

Conclusion

The March of Reason: What Was Hidden in Our Genes

29

James R. Flynn

This chapter has eight parts: pessimism about human genetic potential in the wake of the theory of evolution, cognitive progress in the twentieth century as a historical refutation, how this was captured by massive IQ gains, a paradox posed by the existence of IQ tests, a paradox posed by twin studies, moral progress, reason and morality, and prospects for future progress.

Despite our genes, social progress has enhanced rationality and morality. The industrial revolution had subtle effects on both, and we enter an era in which the measurement of intelligence (IQ tests) raised questions about whether cognitive gains were equivalent to "intelligence" gains. A division of labor solves this problem. The measurement of intelligence properly refers to assessing individual differences in cognitive skills within groups at a particular time and place. The measurement of cognitive progress properly refers to people altering over time: whether they can use reason to deal with a wider range of problems (including moral problems), which is to say with cognitive history. Twin studies posed a dilemma about the potency of environment to account for cognitive progress. The Dickens/Flynn model shows that it can. The relevant question for humanity is whether cognitive and moral progress will persist over the next century.

Adapted from material in James R. Flynn, *Intelligence and Human Progress: The Story of What was Hidden in our Genes*, Elsevier, 2013

J.R. Flynn (✉)
POLS Department, University of Otago,
Box 56, Dunedin 9001, New Zealand

Darwin and the "Scum Worthy"

Darwin had no concept of a gene as a unit of heredity. However, he believed that all creatures inherited characteristics that separated one species from another and also distinguished individuals from one another within species. He was a thoroughly good man but reflected the prejudices of his day regarding the inherited "weaknesses" of people at the bottom of the social scale.

Darwin (1871, p. 510) lamented that physicians prolong the lives of everyone and as a result ". . . weak members of civilized societies propagate their kind. No-one will doubt that this must be highly injurious to the race of man." The man who also independently discovered the theory of natural selection, Alfred Russell Wallace, records a conversation (1890, p. 93): Darwin is oppressed by the tendency of "the lower classes" to over-reproduce and characterizes the surplus as children of "the scum." Wallace's memory could be at fault. However, by 1890, Wallace had totally rejected this image of "civilized society." He was adamant that English society was too corrupt and unjust to allow any reasonable determination of who was fit or unfit. He respected Darwin and was unlikely to so describe his views without foundation.

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_29,

The negative image of the lower classes had deep roots. In *Rob Roy* by Sir Walter Scott (1817), the depiction of near-imbecile servants is quite extraordinary: fidelity to their master is their only saving characteristic. A century later, during World War I, Lord Curzon observed British soldiers bathing: "How is it that I have never been informed that the lower orders have such white skins?" (Blythe 1964). A pity the lower orders were useful as servants. Otherwise these strange white-skinned creatures could have been kept in zoos. During the intervention in Russia in 1918, General Graves of Britain informed General Groves of American that he was getting a reputation as a friend of the poor and that "you should know that these people are nothing but swine" (Melton 2001). The lower classes are scum, rabble, riffraff, louts, peasants, and imbecile yokels sucking on straws.

Most intellectuals greeted the spread of education with a ferocious pessimism (Carey 1992). Virginia Woolf and E. M. Forster were both devoted to adult education. Yet, Wolff refers to the self-taught workingman as someone "we all knew" to be egotistic, insistent, raw, striking, and ultimately nauseating. Forster has no sympathy with a clerk whose attempts to educate himself are "hopeless." He is simply inferior, less intelligent, healthy, and loveable, typical of urbanized rural laborers who should be stripped of their education and revert to what they can do well: breed yeomen. D. H. Lawrence, Pound, Yates, H. G. Wells, George Bernard Shaw, T. S. Eliot, Aldous Huxley, Evelyn Waugh, and Graham Green also derided the capacities of the masses. A rare genius may be hidden among them, but the masses will never match the intellectual attainments and social responsibilities of the elite. The common preference for tinned food is considered damning.

Darwin's fear that the scum will multiply and perpetuate themselves is based on the assumption that the scum of one generation have something about them, something that ensures that their children will be the scum of the next generation.

Today we would say that failure is in their genes. Although we would never be impolite enough to use the word "scum," the thesis is very much alive: the notion that the genes of a substantial part of society mean that their IQ and other personal traits, such as resistance to education, welfare dependency, and criminality, are fixed at a particular time and not subject to modification by new social conditions. Charles Murray believes, as most of us do, that Americans in general deserve a valued place in society appreciated by relatives and associates. But he provides a table in which we are told that, other things being equal, a loss of three IQ points over this generation will mean that the number of women chronically dependent on welfare will increase by 7 %, illegitimacy by 8 %, men interned in jail by 12 %, and the number of permanent high school dropouts by nearly 15 % (Herrnstein and Murray 1994).

Those who do not like the term "scum worthy" can substitute "elimination worthy." Surely that is the cash value of "we want to eliminate your genes because you are likely to have children like yourselves." I reject the thesis of "scum today, scum tomorrow." If you have a fixed pool of "scum," and take their IQ at a given time as a badge of their inferiority, then if they multiply from one generation to another, the percentage of scum increases. On the other hand, if the lower classes can be drained of scum from one generation to another, if they are not permanently scum worthy, society may turn low-IQ parents into higher-IQ offspring and even eliminate undesirable personal traits. As evidence: the whole drift of the last century shows that modernity can alter the minds and capacities of people over time.

Cognitive Progress in the Twentieth Century

Let us forget for a moment that IQ tests were ever invented and focus on people, those peculiar beings that exist even when they are not being tested. We will assume that we do have one measure of cognitive skills, the humble Vocabulary test. Moreover, that it has been standardized from time to time on representative samples of the American population ever since 1950. Therefore, we have a criterion as to what percentage of the US population has a certain level of verbal proficiency at any given time, and we can compare how that percentage altered over time. We also have data from various universities concerning what vocabulary level was a prerequisite for successful study, and census data on the occupational profile of the US population.

In 1900, professionals were 3 % of the population. By 1920, they were still only 5 %. They were held in awe because of their cognitive achievements. Even in 1957, when I went to Eastern Kentucky to lecture, I was referred to reverentially as a "PhD man." By the year 2010, 35 % of Americans were in cognitively demanding jobs: 15 % highly paid professionals and another 20 % subprofessionals, that is, lower management or technical staff (Carrie 2012). There is one possible rebuttal: elite jobs are less cognitively demanding today. Medical colleagues tell me that doctors have to know more science today, commerce colleagues tell me mangers have to plan with a wider range of knowledge, and economics colleagues tell me that today's merchant bankers are virtuosos of cognitive complexity. University academics today sometimes give coherent lectures and do research; university technicians are infinitely more knowledgeable than in the past.

The prerequisite for obtaining most of these jobs is a university degree. Scores on the WAIS (Wechsler Adult Intelligence Scale) Vocabulary subtest can be equated with scores on the Scholastic Aptitude Test Reading Test (Rodrigo de la Jara 2012). This is cheating a bit in that the SAT is an offspring of IQ testing. Nonetheless, the equation tells us what percentage of the US population is viable at leading US universities. The universities will not reveal their minimum score, but there is data for the score that isolates the bottom 25 % of their students (Grove 2012). The average American (50th percentile) is viable at universities such as Corcoran Art & Design, Michigan State, Louisiana Tech, Nevada-Las Vegas, and Fairleigh Dickson (Flynn 2013b). No university flunks as many as 25 %, and therefore, it is realistic to put the vocabulary threshold at a bit below the average, say at the 37th percentile. If you used an IQ metric, that score would be only 5 points below average performance.

Let us go back 50 years to 1960. Jensen (1980) asserts that the average high school graduate was at the 75th percentile and they had only a 50/50 chance of graduating from university. It may be said that elite jobs require a graduate degree. Jensen's data assume that the average candidate in such a degree program was at the 95th percentile and that the minimum standard was about the 88th. Our data show that today the average is the 85th percentile and the minimum standard is the 58th. So in 50 years, we have gone from the top 15 % eligible to get elite credentials to the top 42 %. If the latter seems unrealistic, recall that the top 35 % of Americans hold those jobs today.

Once again, the objection can be put that the universities have set standards below what a university education should require. Well if that is true, how can their graduates do jobs that are cognitively demanding, indeed more cognitively demanding than they were 50 or 100 years ago? The standards of the universities pass what we call the test of external validity. In any event, the brute fact that the masses today fill a huge number of elite jobs falsifies the pessimism current in 1900. The genetic limitations on their rationality did not forbid the social roles once thought the exclusive property of the aristocracy.

And what about altered behavior? WAIS vocabulary gains over time show that adult Americans gained the equivalent of 17 IQ points of active vocabulary in the second half of the twentieth century (Flynn 2013b). This was thanks to the tertiary education revolution. If that gain is projected back to 1900, before the secondary school revolution took place, they made a total gain of 34 points. This is 2.27 standard deviations above the mean and puts them at the 98th percentile of the Americans of 1900. The professionals of 1900 were the upper 3 %. Who would have thought that the average person with an average education could replicate the speech typical of professionals a century ago?

There is additional historical data that attest as to how our minds have altered since 1900. When Luria (1976) interviewed peasants in Russia in the 1920s, he found that preindustrial people had certain cognitive traits in common.

First, they did not classify. When he asked what a fish and crow had in common, they would not say that they were animals. One flies, one swims, you can eat one and not the other. They should not be lumped together because as objects in the concrete world, we use them differently. If you asked someone in 1900 what a rabbit and dog had in common, you use dogs to hunt rabbits. The fact that they were mammals was too incidental to be worthy of notice. Second, they did not take the hypothetical seriously. When asked whether granted that there were no camels in Germany, would there be camels in German cities, they said there must be camels there if the city were large enough. Third, when he asked them to reason about abstractions such as "wherever there is snow bears are white, there is snow at the North Pole, what color are the bears," they stayed firmly rooted in their experience of the concrete world. They had never seen anything but brown bears. But they might believe a reliable witness that came from the North Pole. In frustration they asked Luria how they could solve problems that were not *real* problems.

Today we all know that we do these three tasks readily. We use classification as a means of ordering the world as a prerequisite to understanding it, for example, mammal versus reptile or primate versus non-primate. We take the hypothetical seriously, for example, if medium-sized stars eventually expand into red giants, our sun will do so and destroy the earth. We use logic to order universal assertions, for example, when light behaves both as if it were a particle and a wave, you cannot classify it as one or the other. I call these cognitive traits new "habits of mind."

They are clearly prerequisites for higher education and, as Carmi Schooler (1998) has shown, they allow one to perform the tasks of cognitively demanding jobs. These new habits of mind became so essential that they affected how we educate our children. In 1900, our schools were still firmly rooted in facts about the concrete world. Then they began to teach something new. Genovese compared the exams the state of Ohio gave to 14-year-old schoolchildren between 1902 and 1913 and between 1997 and 1999. The former tested for in-depth knowledge of culturally valued information; the latter expected only superficial knowledge of such information and tested for understanding complex relationships between concepts. Genovese (2002, p. 101) concludes: "These findings suggest that there have been substantial changes in the cognitive skills valued by Ohio educators over the course of the 20th century."

The history of the twentieth century is a story of cognitive progress. The word "progress" is value laden so I will define it by using a hypothetical: if we grant that an expanded vocabulary and our new habits of mind are necessary to comprehend the universe and our own behavior and the modern world, they constitute progress. However, thus far, except for vocabulary, we have no measure of the degree of cognitive progress.

Massive IQ Gains over Time

You measure something when society decides it is valuable enough to measure. When people started to work at dawn and stopped at dusk, what was the need for a personal timepiece? But when the industrial revolution required people to get to work on time, we invented the factory whistle, the clock on the mantle, and the wristwatch. When people inherited their jobs as they did their names, what was the need for an IQ test? But when the industrial revolution required a more educated work force, we invented a measure of who could profit from education, who could progress farthest, and who could become the elite of the modern world. In 1905, Alfred Binet invented the IQ test. French school children told him that something new was worth measuring.

It appears that shortly after a nation embarks on the industrial revolution, IQs begin to rise. Thanks to birth date data (scores rising as the subject's date of birth rises from the past to the present), we know that Britain has made massive IQ gains since 1872. There are data from about 30 nations all over the world, and at their peak, gains run at the rate of at least 0.3 IQ points per year on Stanford-Binet and Wechsler tests, higher on tests like Raven's Progressive Matrices.

Over the last century, IQ gains in Britain and America amounted to at least 30 IQ points.

Scored against today's norms, our ancestors had a mean IQ of 70, the borderline for mental retardation. They were not retarded, of course. Their intelligence was pragmatic: it was focused on how to make use of the concrete world for their own advantage. They lacked our "scientific spectacles," that is, the new habits of mind, the formal education that tutors the mind in logical analysis, and the consequent broad range of vocabulary and general information. The mind of 1900 that is never exposed to such advantages is a far cry from a mind that cannot take advantage of them when exposed (Flynn 2013b).

Thus far, I have emphasized mass education. In fact, causality operated on three levels. The ultimate cause is the industrial revolution or modernity. The intermediate causes are the industrial revolution's by-products, not just enhanced schooling, but a host of other factors. Better standards of living nourish better brains. Family size drops so that adults and their speech dominate the home's vocabulary and modern parenting appears (hothouse parenting or encouraging the child's potential for education). People's professions exercise their minds rather than asking for physically demanding repetitive work. Leisure allows cognitively demanding activity rather than mere recuperation from work. The world developed a new visual environment so that abstract images dominate our minds and we can "picture" the world and its possibilities rather than merely describe it. The proximate causes are the minds people take with them into the test room so they can answer more items correctly, not simply their new "habits of mind" (classification, logical analysis of abstractions) but also vocabularies, general information, and visual awareness.

IQ gains are not eternal. Sooner or later, the intermediate causes gradually lose potency. Education is widespread and adequate, family size can go no lower, and leisure is as packed with as many cognitively demanding pursuits and images as anyone can tolerate; even featherbedding can produce no more elite jobs, so the triggers of massive IQ gains stop.

America and Britain show IQ gains over 100 years or more and are still advancing. However, more progressive societies such as Scandinavia and the Netherlands appear to have emerged from the IQ gains period. The period for some nations may fall well short of 100 years. China and Japan and Korea industrialized much later than America and Britain, and their rate of social change has been dramatic. Rapid social change has put their rate of gain well above the US-British rate, but the price they pay may be a shorter cycle. Developing nations that have really begun to develop, Argentina, Brazil, Turkey, and Kenya, are just entering their massive gains phase. Much of the world is still in the doldrums. The next century will be interesting. Developing nations are some 10–30 IQ points behind the developed world. But there is strong evidence that those favored by economic development (Latin America in particular) will catch the developed world within 40 years. Much of the developed world is likely to remain in the doldrums or regress under the impact of climate change (Flynn 2013a).

Given what modernity has done to the human mind, what kinds of IQ tests or subtests would we expect to be most affected? Every nation in its IQ gains phase has made enormous gains on Raven's Progressive Matrices. Indeed, the best estimate (remember we have birth date data from 1872) is a total gain of over 50 points in 100 years. This test above all measures the use of logic on abstractions (matrices patterns) removed from the concrete word. In essence it is a kind of analogies test.

Fox and Mitchum (2012) have analyzed just what has allowed each generation to do better on Raven's than the preceding generation. One hundred years ago, Americans could do simple analogies grounded in the concrete world: Domestic cats are to wild cats as dogs are to what? (Wolves.) This would do them no good on the kind of analogies found on Raven's. But by 1961, they could handle two squares followed by a triangle implying two circles followed by what? (A semicircle: just as a triangle is half of a square, so a semicircle is half of a circle.) By 2006, they could handle two circles followed by a semicircle implying two sixteen's followed by what? (Eight: you have to see the relationship despite the transition from shapes to numbers.) Note how each step takes us further from the concrete world toward using logic on abstractions, eventually abstractions whose very identity shifts. Who can imagine the average person in 1900 able to do all of that? Is it any wonder that we get much higher scores on Raven's?

We have referred to Wechsler gains. Where have these been the largest? They have been the largest, first, on the Similarities subtest that forces you to classify; second, on Analytic subtests that force you to use logic to devise how blocks or objects can make certain designs; third, on the Pictorial subtests which ask you to find the missing piece of a picture or use pictures to tell a story; and fourth, on the Vocabulary subtest where adults made large gains thanks to more and more education. In recent years, children have had no more additional schooling, and their vocabulary gains have been modest (Flynn 2013b)

In sum, the historical evidence and the pattern of IQ gains match. The enormous score gains are a measure of the enhanced cognitive traits that distinguish the modern mind from the minds of our immediate ancestors.

But Are They Intelligence Gains?

The argument thus far rests on two syllogisms. First, the cognitive demands of elite jobs and education are greater than 100 years ago; many more people can meet those demands; therefore, there has been cognitive progress. Second, classifying the world, using hypotheticals, and using logic to render generalizations consistent are more cognitively complex than simply taking the concrete world as a given; far more people can do the former; therefore, there has been cognitive progress.

Even if no IQ tests existed, any aware person can see that his or her mind differs profoundly from the American mind in 1900. However, IQ tests do exist and their record of gains over time offers a bonus: we can actually measure the degree of cognitive progress modernity confers. It would be odd if this were not the case. IQ tests were *designed* to measure the traits that were enhanced: logical analysis, analogies, classification, pictorial awareness, vocabulary, and general information. Given the evidence, it would seem that those who hold IQ tests in the highest esteem would be the first to concede cognitive progress. In fact, the opposite was true.

Those who follow the late Arthur Jensen deny that IQ gains over time are intelligence gains. Jensen (1998) called them "hollow," lacking real-world cognitive significance because they could not pass what he called the "method of correlated vectors." Here we must introduce *g*, often called the general intelligence factor. There is nothing mysterious about *g*. Something similar exists in many areas. Some people have "musical *g*": whatever instrument they pick up, they learn quickly. Others have "athletic *g*": they shine at all sports. There is a strong tendency for the same people to score above or below average on all of the 10 or 11 Wechsler subtests, no matter whether they test for vocabulary, general information, mental arithmetic, solving three dimensional jigsaw puzzles, or discerning logical relations conveyed by a matrix.

Factor analysis measures the strength of the tendency of various subtests to be intercorrelated. You can then go back to the subtests and calculate a hierarchy as to how much performance on each of them predicts general performance across the whole set of subtests. This is their *g* loading. The best predictor is usually (not always) your performance on the Vocabulary subtest. Now you can rank the ten subtests from those that have the greatest "*g* loading" down to those that have the least. Jensen then ranked the same tests from those whose score gains over time were the greatest down to those whose score gains were least. If the subtest gains have a negative correlation with the *g* loadings, and there is a mild tendency in that direction, you conclude that IQ gains are not really intelligence gains. This assumes, of course, that it is legitimate to identify intelligence with *g*.

We can see why Jensen thought the identification appropriate. The impressive thing about the *g* loadings of subtests is that they rise with the degree of cognitive complexity of the task the subtests measure. As Jensen often pointed out, the *g* loading of digit span forward, a simple task of repeating a series of random numbers in the order in which they are read out, has a low *g* loading. Digit span backward, a more complex task of saying numbers in reverse of the order in which they are read out, has a much higher *g* loading. Speed of shoe tying would have a *g* loading of close to zero. Most of us feel that the more cognitively complex a task, the more it measures intelligence.

And yet, Jensen's demand leads to a paradox. People over time have made huge gains on subtests every one of which poses problems of cognitive complexity. Yet all of these gains are to be dismissed because the gains did not privilege tasks according to the *magnitud*e of their cognitive complexity. Imagine we added shoe tying as an 11th subtest and for some reason, perhaps enhanced dexterity over time, people make by far the largest gains on it. This would virtually guarantee a negative correlation between IQ gains over time and cognitive complexity. The solution to this paradox lies in whether the demand for a hierarchy of cognitive complexity is a legitimate demand both for measuring intelligence and measuring cognitive progress. I will argue that it is appropriate for the first but not the second.

Take two people at a given place and time sharing the same cognitive environment (two brothers in the same home). If one accesses that environment better than the other, it makes sense to say he has the better mind. Moreover, he is likely to outstrip his brother in accord with cognitive complexity. The less able brother will not fall far behind for simple cognitive skills, but he is more likely to fall behind for complex ones. After all, they live in a shared cognitive environment: both are subject to hothouse parenting, both will enjoy much the same amount of schooling, both have modern habits of mind, and so forth. I have complicated views about "intelligence," but in this context, I am willing to call the difference between their IQs, particularly when weighted for *g* loadings, an intelligence difference. Van Bloois et al. (2009) have done an excellent study showing that the gifted, the average, and the mentally retarded differ on Wechsler IQ subtests in accord with *g* loadings.

Society, on the other hand, is quite different. It does not administer a gigantic IQ test, rank mental skills in order of their degree of cognitive complexity, and then decide to enhance them going from top to bottom. It actually responds to real-world social priorities. If it needs mass education and people to fill chattering jobs (law, teaching, counseling), it will enhance vocabulary. If it needs executives to do lateral thinking, it will encourage using the hypothetical. If it needs a wider range of information to cope with a more complex modern world, it will enhance general information. In a post-sputnik era, if it wants more people adept at mathematics, it will push arithmetical skills – and if it does not know how to improve them, gains will be slight despite their high *g* loading. If the fact everyone has their own car enhances the need for navigational skills, map-reading skills will go up despite its low *g* loading.

In other words, when society shifts its priorities for what mental skills are needed over time, it cares absolutely nothing for sheer cognitive complexity. It makes no sense at all to advise it to respect a *g* hierarchy. To demand this is to confuse society with a giant brain.

Individuals have brains. Genes influence their overall quality; they probably give some people an optimum blood supply to the brain and an optimum dopamine spraying system. Certain neurons spray dopamine, which strengthen the neural connections in the brain with use, rather like having a good sprinkling system for your lawn. When we compare generations over time, we are not comparing one gigantic brain to another, both operating in a common cognitive environment, which the later brain accesses more efficiently. We are comparing two complex social systems whose altering cognitive priorities create radically different cognitive environments. If the environment has become more cognitively demanding, there is cognitive progress. No one's brain is any better at conception, and no one is more intelligent in the sense of adapting better to a common environment. The rise in the average IQ compares two cognitive environments, not individual differences. Various mental abilities alter autonomously, that is, without regard to *g* or relative cognitive complexity.

It may be asked: When cognitive skills are enhanced autonomously, does this have any real-world significance? An accumulating number of studies show that the answer is yes.

Coyle and Pillow (2008) show that when you deduct *g* from performance on the SAT, the scores still predict university grades. Ritchie, Bates, and Deary (under review) show that the effects of education are beneficial even though they are not mediated by *g* but consist of direct links to specific subtests. Woodley et al. (Woodley 2012a, b) show that education actually tends to promote diversified cognitive skills and that IQ gains over time (which of course do not correlate with *g*) parallel and predict growth in GDP per capita. Woodley (in press) concludes that autonomous skills allow one to adapt cognitively to modernity and thereby promote a better life. Armstrong and Woodley (under review) show that modernity in general encourages greater sensitivity to a whole range of rules, ones that operate independently in a complex web of social situations, rather than collectively as assumed by *g*. Finally, for the specialists, Fox and Mitchum (2013) show that enhanced performance on Raven's is not due to hollow skills (like test sophistication) but to real-world cognitive advance, even though the skills enhanced are not correlated with *g* and are *not factor invariant*.

So now a simple division of labor has solved the paradox. We will restrict the use of *g* hierarchies to assessing *individual differences* between people sharing a relatively homogenous cognitive environment at a given place and time. And we will eschew *g* when assessing what *generational differences* over time occur as people's minds alter thanks to altered social priorities. Honor will have been preserved for all. We will never contaminate *g* by calling cognitive progress "intelligence gains," and we never dismiss cognitive gains by demanding that they be *g* gains. However, the two are kissing kin: both have to do with enhanced ability to solve cognitively complex problems, one by individuals in pecking order, the other by generations helter-skelter.

I should add that I do not mean to imply that the concept of *g* is trivial. The fact that *g* loadings correlate with cognitive complexity is illuminating. We must rely on our intuition to establish that the two correlate at all but specific cases are convincing. There is the fact that digit span forward (simple task of rote memory) has a low *g* loading, while digit span backwards (more complex) has a higher *g* loading. Making a soufflé has a higher *g* loading than scrambling eggs. Once we accept the relationship, it is illuminating. Vocabulary (assuming equal opportunity) ranks minds for the cognitive complexity of the concepts they can absorb. Arithmetic ranks minds for how well they can plan a numerical strategy and carry it out mentally (without pen and paper). Which of the two involves more cognitive complexity? Vocabulary has the higher *g* loading – fascinating.

The hierarchy of *g* loadings correlates with the degree to which inbreeding (negatively) influences subtest performance. This shows that those areas of the brain that do cognitively complex mental tasks have a genetic substratum more fragile than those areas that do less complex tasks. They are more subject to damage by the pairing of undesirable genes during sexual reproduction. This is what inbreeding enhances. We have a significant contribution to our knowledge of brain physiology.

The Tale of the Twins

Twin studies (and other kinship studies) challenge not the significance of cognitive progress but whether a coherent account can be given of the causes of IQ gains.

Take identical twins that were separated at birth and raised by different families. If they grow up with identical IQs, the inference is that identical genes trump dissimilar and enfeebled environments. If they grow up with IQs no more alike than the rest of us, dissimilar environment has trumped identical and enfeebled genes. The result: they are far more alike for IQ than randomly selected individuals. By adulthood, all kinship studies show that family environment has faded away to zero. Adult IQs differ only to the degree that chance events might cause them to differ (one is dropped on his head and the other was not). It is hard to see how chance events could differ between the generations and cause massive IQ gains over time.

Thus, environment is too feeble to have much influence on IQ. Yet, massive IQ gains over time occur whose causes appear to be overwhelmingly environmental. We have a new paradox: How can solid evidence show that environment is both feeble (kinship studies) and potent (IQ gains) at the same time? The Dickens/Flynn model solves this paradox (Dickens and Flynn 2001a, b, 2002).

Let us see what happens to children that are genetically identical but grow up in different families. I will use basketball as an example. Joe and Jerry are identical twins separated at birth so that one is raised in Muncie Indiana and the other in Terre Haut. Thanks to their identical genes both will be four inches taller and a bit quicker than average (faster reflex arc). Indiana is a basketball mad state, and at the start of school, both boys get picked to play sandlot basketball more often than other kids. This is the beginning of matching above average genes with an above average environment. Moreover there is reciprocal causation between their skills and their environment: better skills mean a better environment, which upgrades their skills, which means an even better environment, and so forth, essentially a feedback mechanism. The Dickens/Flynn model calls this the *individual multiplier*.

Next they make their grade school teams, which upgrade their skills further, and they both make their high school teams and get professional coaching. These separated twins will end up with highly similar basketball skills, but why? Not merely because of their identical genes but also because of their highly similar basketball histories. In the kinship studies, genes get all the credit and basketball environment gets nothing. But this is a misinterpretation. It pretends that environment is feeble, when in fact their genes have co-opted something as potent as more play, team play, and professional coaching. Potent environment is disguised simply because it is matched with identical genes.

Now let us shift to factors that affect the collective basketball environment over time. The genes of people in general are essentially static over a few years, so now basketball environment is cut loose from genes and emerges in all its potency. After World War II, TV was invented and the close-ups of basketball were exciting and popularized the sport. Far more people participated and this raised the skill level. Indeed the rising average performance became a causal factor in its own right and a new feedback mechanism was born, which we call the *social multiplier*.

To be above average, it was initially good enough to shoot and pass well. Then ambidextrous people began to pass with either hand and find more open players, and the rising mean forced everyone who wanted to keep up to do the same. Then people began to shoot with either hand and get more opportunity to score baskets because they could go around a guard on either the right or left side. Almost overnight basketball was transformed from the stodgy sport of 1950 to the incredibly fluid and graceful sport that took root in the 1960s.

The comparative potency of genes or environment depended on whose hand was on the throttle of a multiplier. Comparing individuals within a cohort, genes co-opted environment and genes seemed omnipotent, thanks to the individual multiplier. Comparing generations over time, evolving environment broke free to raise the performance in basketball to new heights, thanks to the social multiplier.

I take it the analogy is obvious. Identical twins in separated environments may have genes that set them above (or below) the average person for cognitive ability. If above, what are small genetic differences at birth become potent because they co-opt matching and superior cognitive environments: more attentive teachers, superior peer interaction, honor streams, and better high schools and universities, factors hardly rendered impotent simply because they are co-opted by genes. Over time, things are different. Increasing the years of schooling from six to twelve to more than twelve (university) really does do something to enhance the cognitive abilities of the whole society. The mere fact that genetic differences tend to determine how many years of school a person gets at a given time does not weaken the potency of additional years of schooling over time.

Just as the near identical scores of separated identical twins do not rob environment of its potency, the huge environmentally induced IQ gains over time do not rob genes of their potency. They are both potent enough to do their jobs, explaining individual differences versus explaining group differences over time.

The multipliers also solve a problem that baffled the psychological community. If environment is weak within groups, then to explain huge environmental effects between generations over time, you have to invent a factor X: a mysterious environmental factor that operated exclusively between groups or generations.

We now see that much the same factors are operating within and between groups. Within groups, individuals are distinguished by factors like better families, teachers, peers, universities, and jobs. These factors are made to seem feeble because the individual multiplier correlates them with genetic differences, and twin studies show them as having little impact beyond what genetic differences would dictate. Between groups, the two generations are also distinguished by factors like better parenting, more schooling, and more cognitively demanding jobs. But thanks to the social multiplier, they have huge effects simply as environmental variables. They operate free of genes because there are no real genetic differences between the generations that they *could* be correlated with.

In sum, the "weakness" of an environmental factor within groups is a mere appearance and does not translate into weakness between groups. Much the same environmental factors operate both within and between groups and no mysterious factor X is necessary. The factors that separate generations do not necessarily, of course, apply to ethnic groups. Black subculture digs a gulf between black and white that is peculiar to those two groups (Flynn 2008).

The London Mob

Many members (not all) of the Victorian elite were pessimistic about moral progress. This was partially based on "the London mob." In 1780, the House of Commons refused to debate a petition against granting Catholics toleration. The poor, criminals, and prostitutes rioted with hundreds killed and some hanged (German and Rees 2012). Although this was their last great riot, the practice never disappeared and the respectable classes' image of mass violence persisted. Thirty books that appear after 1840 express apprehension: "Now it is the general complaint of the taverns, the coffee-houses, the shopkeepers and others, that their customers are afraid when it is dark to come to their houses and shops for fear that their hats and wigs should be snitched from their heads or their swords taken from their sides, or that they may be blinded, knocked down, cut or stabbed; nay, the coaches cannot secure them, but they are likewise cut and robbed in the public streets, etc." (Shoemaker 2004, p. 162).

When we turn to at our genetic inheritance, the pessimism of the elite about moral progress may seem to haves some substance. Our nearest primate relatives suggest that over much of human evolution, males and females were subject to different selective pressures.

Males competed for access to females by either violent combat or aggressive displays that intimidated rivals. Since aggressive males fathered the most offspring, their genes became dominant. Females perpetuated their genes to the extent that they raised their children to maturity, so that their children could reproduce. A bond with a male helpmate was advantageous. Therefore, genes for whatever helped domesticate males were positively selected. These proclivities prepared the way for the emergence of traits that statistically differentiate the genders. It is politically incorrect to assert that women are cleaner, more attentive to physical appearance, more skilled at arts that make home life attractive, and more likely to use charm rather than (overtly) aggressively behavior to attract the opposite sex. I will reply on those of both sexes who see through their eyes and not their ideologies.

However, there is reason to believe that our genes have altered. Hallpike (2008) points out that male aggression began to pay decreased procreative dividends in the simplest Homo sapiens societies, the hunter-gatherer societies universal until about 10,000 years ago. The simple societies that survive today show that the collective action of other males can eliminate or expel an overly aggressive male. The best hunter is expected to share his kill (spoilage makes most of it worthless to him). About 10,000 years ago, human beings started living in larger communities, which were functional only if aggression was restricted by rules. Just as people domesticated animals like dogs and cats, people began to domesticate themselves. Just as domesticated animals were selected for self-control of their aggressive behavior, not to be directed at their masters but to be governed by rules the master set, so people were domesticated by genetic selection for self-control and rule-bound behavior (Wilson 1991; Leach 2003).

Steven Pinker (2011) comes onstage at this point. The growth of larger cities and nations increased the range of people that the inhabitants were "trained" to forgo aggression against. Trade within and between nations was an important factor: you want to preserve a lucrative customer, not kill him and confiscate his property. To put the point in evolutionary terms, assume that over 1,000 generations law-abiding citizens have outreproduced those predisposed to violence. If so, human genes were selected so that we find it easier today to live together without physical aggression. This is plausible but unproven: evidence will follow.

Over the last 1,000 years, there has been another domesticating trend. Males are responsible for most acts of violence. Female domestication of males is signaled by the fact that males are violent primarily between puberty and sometime in the 20, after which they are pacified by the responsibilities of marriage and child rearing (Pinker 2011). As civilization developed, male competition for women focused less on violence and more on money, status, and amiability (a nice guy).

Over the last few generations, some women gained power to pacify males because of trends that empower them in the home: the ability to find employment so that they need not be totally dependent on males to support themselves and their children, the presumption that both sexes will contribute to home maintenance and child rearing, the fact that division of property and child support means that a male cannot evade responsibilities through divorce, and legal sanctions against domestic violence. Some societies lag. Many Sunni Muslims allow a man to divorce his wife by saying "I divorce thee" three times. The husband is not responsible for the wife's expenses (but responsible for the maintenance of children until they are weaned).

Middle Eastern men have become aware of what they face. Virk (2012) says that historically men have tended to be free spirited, adventurous, and wild. He describes five stages of domestication: courtship (a man wears clothes and uses perfumes agreeable to women and affects an interest in culture), declaration (he must express love rather than compliments), employment (he must get a job so she can hold him in esteem), home ownership (she tortures him with an account of how their landlord tried to take sexual liberties and suggests that rather than killing the landlord, the obvious remedy, he buy a home of their own), and parenthood (she begins to call him childish names such as "baby" and shifts child care onto him). His complete domestication is signaled when "they go to market with a baby hanger on the husband's back and a patent little handbag in the wife's arm." This description reveals, I fear, a determination to fight in the trenches.

Now we turn to the evidence. If domestication has occurred, violence should have declined. The seventeenth and eighteenth centuries shift away from cruelty: fewer amusements like roasting a cat alive or men competing to batter a pig to death with clubs. The last heretic was tortured; the last witch burned in Europe. Slavery had existed for thousands of years. In the nineteenth century, "an overwhelming majority of Westerners came to feel that slavery was *wrong*." Dueling is gone and clan feuds, gang wars, and lynching are nothing compared to a century ago. We no longer glorify winning the West by killing Indians. Slaughtering people on the highways so we can enjoy drink, the intoxication of speed, or the manly desire to use a car as a tool of combat is questioned (Brinton 1959).

Pinker (2011) adds quantified evidence. As for violent death from war, hunter-gatherers (14000 BC to 1770 BC) get a rate of 15 per hundred deaths. Beginning with the early cities and empires of recorded history, the rate falls to 3–5%.

Perhaps a better measure is the chance the average person has of dying from violence in a given year. In thirteenth-century England, the homicide rate was over 20 per hundred thousand per year. From the sixteenth to the twentieth century, the rate steadily dropped down to less than one throughout Europe. America shows about 5 people per 100,000. The "far north" of America (New England west to Oregon and Washington) is as safe as Europe but homicides escalate as you go south.

Pinker calls the period after 1946 the long peace. There are civil wars and great powers bully minor ones. However, no great power has engaged in direct combat with another. Until recently, the expectation that great powers would fight one another was normal. The Hapsburgs, Spain, France, England, Russia, Germany, America, Italy, Turkey, Japan, China, the Netherlands, and Sweden all did so. Finally, since 1989, there has been the new peace. Civil wars, genocides, repression by autocratic governments, and terrorist attacks have all declined. We have a new commandment: no wars shall be fought to annex territory across national boundaries.

Reason and Morality

We return to our main theme: the consequences of the march of reason. Georg W. Oesterdiekhoff (2009) has traced the effect on ethics running from magic to religion to a scientific knowledge of reality. Magical and religious beliefs produced much immoral behavior in the past: human sacrifice to feed the gods (the Aztecs killed about 1.2 million), the burning of witches, and the horrors of the Inquisition. In the Old Testament, God instructs his chosen people to slaughter animals. When Aaron's two sons do so using the wrong kind of incense, he burns them alive. God is aware that a captured woman may not be in the mood for sex having seen her husband and children slain. God advises the Israelites to shave her head, pare her nails, and imprison her until she sees the wisdom of being raped (Pinker 2011).

Are we fully aware of what it was like to live everyday life surrounded by superstition? In many tribal societies, every natural death was a murder and innumerable innocent people were executed. It is horrible to contemplate that some of them thought that they were guilty: What if they had wished the person dead or had dreamed about their death? When murders occurred, using divination to establish the guilty party was counterproductive. Many murderers walked free ready to kill again.

The personification of animals was inherited from tribal society. From the thirteenth to the eighteenth centuries, animals thought complicit in murder, assault, plague, or bestiality were tried and executed throughout Europe. They included pigs, horses, bulls, cows, sheep, rats, beetles, and insects. Some were clearly wronged: in 1474, a rooster was prosecuted for laying an egg fathered by Satan. Some were exonerated: lawyers won famous victories representing rats and beetles (Evans 1906).

Tertullian extolled holy ignorance: "We have no need of curiosity after Jesus Christ, nor of research after the Gospel." Fortunately, Europe did not heed him.

By 1900, the new scientific ethos had blind faith on the defensive. However, it did little to banish the secular demons of racism and nationalism that culminated in the horrors of World War II and the holocaust. We were still like domesticated animals. We had selected ourselves to resist violence within groups but not between groups: we were happy to coerce "inferior" races or kill traditional enemies. Yet, over the last 70 years, these demons have been on the defensive.

During the twentieth century, we made an enormous leap forward in adopting a new moral notation. The story of mathematics is the story of improved notation. Greek symbols were so cumbersome that it took the genius of Archimedes to represent large numbers. Roman numerals were an advance but contemplate the task of dividing MDCCCVIII by IV: the answer is 452 (1808 divided by 4). The modern mind has a new way of stating moral maxims. Remember Luria. We are now ready to generalize, challenge generalizations by using hypotheticals, and demand that they be logically consistent with one another.

In 1955, Martin Luther King began the Montgomery bus boycott. My brother and I argued with our father: “What if you woke up tomorrow and had turned black?” Reply: “That is the dumbest thing you have ever said, who do you know that turned black overnight?” He simply would not take the hypothetical seriously. My *Beyond Patriotism* (2012) diagnoses the retreat from nationalism since Vietnam. “What if your home was hit by a drone because someone nearby was sheltering a Taliban?” Or “If a war killed foreigners to save 3,000 Americans, where would you fall off the boat: at 10,000 or 100,000 or one million?” The answer tends to divide youth from age (the latter: “their government protects them and our government protects us”). Inherited maxims can be very cruel. Islamic fathers shock the world when they kill a daughter because she has been raped. We would ask: “What if you had been knocked unconscious and sodomized?’ He is unmoved. He sees moral maxims as concrete things, no more subject to logic than any other concrete thing like a stone or tree.

Today we worry about the “collateral damage” of killing foreigners in Afghanistan and Pakistan. No military commander uses language like “bombing the Vietnamese back to the stone age.” In 1914, Thomas Mann says he had long felt the need of a war to subordinate materialism to “German *Kultur.*” Rilke called the war the resurrection of “the God of hosts.” Max Weber gushed “this war is great and *wunderbar.*” Even the saintly Martin Buber lost his mind: “I know personally that Belgian women amused themselves by putting out the eyes of wounded German soldiers and forcing buttons ripped from their uniforms into the empty eye sockets” (Elon 2003). List the men of letters who would talk like that today.

There have always been people who were antiracist and anti-nationalistic and subscribed to something like the golden rule: the innocent should not suffer; put yourself in their place. But for the overwhelming majority of humanity the golden rule was merely one of a host of inherited maxims: blacks should know their place, my country right or wrong, the obligations of honor (kill my daughter), and the “rights” of the individual (own my own gun). They might find some of these inherited things more attractive than others, but that did not mean they had the new habits of mind that upgrade moral debate. Valuing your possessions is not the same as testing generalizations against logic.

The UNODC (2013) has found that national differences in homicide rates correlate better with intelligence (measured by IQ and school achievement) than with years in school, GDP, less corruption, and greater freedom and democracy. Hodson and Busseri (2012) found that low IQ in childhood predicted racism, homophobia, and membership in groups inclusive of hypernationalism. Still, too many variables are correlated with IQ to use this evidence to single out patterns of moral reasoning as a potent variable. I can only appeal to the historical record.

Progress at Risk

Will science and rationality spread to embrace the world? Unfortunately, something that has pacified humanity undermines steps to deal with the most important threat we face, and some actors ignore the most important rule for maintaining peace.

In the past, the more nations that enjoyed economic progress and engaged in international trade the better. Today, the momentum of economic progress promises to make cutting carbon levels in the atmosphere impossible. There has never been a time in the earth’s history when the carbon content of the atmosphere has been above 1,000 ppm (parts CO_2 per million) and when the polar ice caps still existed. A race goes on between how much carbon dioxide we emit per unit of economic output, which is diminishing by 1.3% per year, and how fast economic growth escalates, which is about 3.45 % per year. Projections: we will pass 500 ppm by 2050 and the no polar ice caps value of 1,000 ppm soon after 2100. To win this race, the rate of economic growth would have to fall. People in the developed world would have to lower their standard of living. The present trend of raising people in the developing world out of poverty would come to a tragic end (Flynn 2013b).

The Kyoto talks are going nowhere. What American President is going to accept targets that would have him face reelection on a platform of less prosperity? What Chinese leader is going to tell his rural poor that they are going to stay poor? How can we stop temperature rise without cutting the growth rate that is the only hope of the world's poor? Stephen Salter of Britain has proposed by far the least dangerous method. At a cost negligible compared to the costs of climate change, a fleet of ships would send sea spray upward to whiten the clouds and reflect away the sun's heat. This would actually lower the earth's temperatures, and in the meantime, we might develop clean power: using lasers or plasma to achieve hydrogen fusion.

The territorial commandment that "no one uses force to annex territory" is fundamental to banning war over the twenty-first century. The Middle East is volatile because of the antagonism between Sunni Muslims and Shiite Muslims. There is also an ambiguity that creates a far more dangerous situation. Many regard the 1967 border between Israel and the occupied territories on the West bank of the Jordan as the potential border between Israel and a Palestinian state. Thus, even moderate Arab opinion sees Israeli expansion of settlements in that area as a violation of the territorial commandment. This gives camouflage to extremists who preach a crusade to eliminate Israel. Terrorist groups harass Israel with all means of sabotage they can command. When they get drones, this may reach intolerable levels.

The statement of the problems we face forbids optimism. Still, whatever happens to us, we can take satisfaction in how far we have come. Living our lives day by day, we take modernity for granted. The very existence of the modern world is astonishing. I refer not to the Internet or the air travel or the organ transplants but to the people. No totalitarian regime created a "new man" but without fanfare impersonal social forces have begun the task. The upper classes were so confident that the masses could never match their intellectual attainments and social responsibilities. They were so confident that the London mob could never be pacified. They were wrong. As Kipling (1996) put it, "For the Colonel's lady and Judy O'Grady are sisters under their skins."

References

Armstrong, E. L., & Woodley, M. A. (under review). The rule-dependence model explains the commonalities between the Flynn effect and IQ gains via retesting.

Blythe, D. (1964). *The age of illusion*. Boston: Houghton Mifflin.

Brinton, C. (1959). *A history of Western morals*. New York: Harcourt.

Carey, J. (1992). *The intellectuals and the masse: Pride and prejudice among the literary intelligentsia, 1880–1939*. London: Faber.

Carrie, A. (2012). Occupation change: 1920–2010. Weldon Cooper Center for Public Service: http://statchatva.org/2012/04/06/occupation-change-1920-2010/ Note: The 1900 census did not use the same system of classification. However, Carrie put 1920 at 5 percent professionals and 1910 was 4 percent: Durand, E. D., & Harris, W. J. (1999), Population 1910: Occupational statistics (United States Bureau of the Census), New York: Norman Ross (Table 14). Therefore, I put 1900 at 3 percent.

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA after removing *g*. *Intelligence, 36*, 719–729.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: Macmillan.

de la Jara, R. (2012). How to estimate your IQ based on your GRE or SAT scores. Google "IQ comparison site".

Dickens, W. T., & Flynn, J. R. (2001a). Great leap forward: A new theory of intelligence. *New Scientist*, 21 April: 44–47.

Dickens, W. T., & Flynn, J. R. (2001b). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*, 346–369.

Dickens, W. T., & Flynn, J. R. (2002). The IQ paradox is still resolved: Reply to Loehlin and Rowe and Rodgers. *Psychological Review, 109*, 764–771.

Elon, A. (2003). *The pity of it all: A portrait of the German-Jewish epoch 1743–1933*. New York: Picador.

Evans, E. P. (1906, 1987). *The criminal prosecution and capital punishment of animals*. London: Faber and Faber.

Flynn, J. R. (2008). *Where have all the liberals gone: Race, class, and ideals in America*. Cambridge, UK: Cambridge University Press.

Flynn, J. R. (2012). *Beyond patriotism: From Truman to Obama*. Exeter: Imprint Academic.

Flynn, J. R. (2013a). The "Flynn effect" and Flynn's paradox. *Intelligence*, 41, 851–857.

Flynn, J. R. (2013b). *Intelligence and human progress: The story of what was hidden in our genes*. London: Elsevier.

Fox, M. C., & Mitchum, A. L. (2012). A knowledge based theory of rising scores on "culture-free" tests. *Journal of Experimental Psychology: General*. doi:10.1037/a0030155.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge based theory of rising scores on “culture-free” tests. *Journal of Experimental Psychology: General, 142*, 979–1000.

Genovese, J. E. (2002). Cognitive skills valued by educators: Historic content analysis of testing in Ohio. *Journal of Educational Research, 96*, 101–114.

German, L., & Rees, J. (2012). *A people’s history of London*. London: Verso.

Grove, A. (2012). College admissions. About.com.guide (search by each state).

Hallpike, C. R. (2008). *How we got here: From bows and arrows to the space age*. Central Milton Keynes: AuthorHouse.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.

Hodson, G., & Busseri, M. A. (2012). Bright minds and dark attitudes. *Psychological Science, 23*, 187–195.

Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. New York: Praeger.

Kipling, R. (1996). *The Ladies*, Stanza VIII, in *Barrack-room ballads* (2nd series).

Leach, H. M. (2003). Human domestication reconsidered. *Current Anthropology, 44*, 349–368.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Melton, C. K. W. (2001). *Between war and peace: Woodrow Wilson and the American expeditionary force in Siberia, 1918–1921*. Macon: Mercer University Press.

Oesterdiekhoff, G. W. (2009). *Mental growth of humankind in history*. Norderstedt: Norderstedt Bod.

Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes*. London: Penguin.

Ritchie, S. J., Bates, T. C., & Deary, I. J. (under review). Does education boost general intelligence (*g*) or specific cognitive abilities?

Schooler, C. (1998). Environmental complexity and the Flynn effect. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 67–79). Washington, DC: American Psychological Association.

Scott, S. W. (1817). *Rob Roy*. Edinburgh: Archibald Constable.

Shoemaker, R. B. (2004). *The London mob: Violence and disorder in eighteenth- century England*. London: Hambledon & London.

United Nations Office on Drugs and Crime. (May 2013). *Homicides 2008.* New York: United Nations Publications.

van Bloois, R. M., Geutjes, L. L., te Nijenhuis, J., & de Pater, I. E. (2009). *g* loadings and their true score correlations with heritability coefficients, giftedness, and mental retardation: three psychometric meta-analyses. Paper presented at the 10th Annual Meeting of the International Society for Intelligence Research, Madrid, Spain, December.

Virk, S. H. (2012). The domestication theory. http://worldpulse.com/node/49936

Wallace, A. R. (1890). Human selection. *Popular Science Monthly, 38*, 90–102.

Wilson, P. J. (1991). *The domestication of the human species*. New Haven: Yale University Press.

Woodley, M. A. (2012a). A life history model of the Lynn-Flynn effect. *Personality and Individual Differences, 53*, 152–156.

Woodley, M. A. (2012b). The social and scientific temporal correlates of genotypic intelligence and the Flynn effect. *Intelligence, 40*, 189–204.

Woodley, M. A., Figueredo, A. J., Ross, K. C., & Brown, S. D. (in press). Four successful tests of the cognitive differentiation-integration effort hypothesis. *Intelligence, 41*, 832–842.

30 Closing Comments: Intelligence and Intelligence Tests – Past, Present, and Future

Jack A. Naglieri and Sam Goldstein

For nearly 3,000 years, philosophers, educators, physicians, scientists, and psychologists have discussed, debated, and written about the apparent singular force of intelligence that has evolved in all species but found its greatest progression in humans. Yet, it has been a little more than 100 years that the concept has dramatically evolved as a powerful phenomenon used to make critical, often life-changing decisions about people. Often lost in the debate over "what is intelligence and how should it be measured" is an appreciation of the evolutionary basis of this concept. In this volume, we have sought to push the historical roots of this concept in ourselves and all species on this planet to the forefront. We argue and debate the role of intelligence in education, yet it appears to us that few of the participants in this often heated controversy have taken the time to understand and appreciate this concept on an evolutionary basis. As the joke goes that our nose evolved to hold our glasses, so too can we apply this blind logic to conclude that intelligence evolved to make academic decisions about children. Intelligence, as the many contributors to this volume can attest, however, is about solving problems in better and effective ways. We have chosen to address this closing chapter not to the thousands of years of discussion about intelligence as a philosophical or scientific concept but rather to the last 100 years and the myth of IQ.

The history of IQ and intelligence testing has been mired in controversy, acrimony, and mistrust leading some to characterize the field in the most negative of terms. For example, Ridley writes "few debates in the history of science have been conducted with such stupidity as the one about intelligence" (200x, p. 77). Ridley's criticism may be considered an overstatement, but it is clear that intelligence as a measurable phenomenon was poorly defined at the outset. The lack of a firm theoretical basis for IQ tests was noted by Pintner (1923) when he wrote that test developers "borrowed from every-day life a vague term implying all-round ability and knowledge" and are still "attempting to define it more sharply and endow it with a stricter scientific connotation" (p. 53). Despite the vague origins of the concept of general intelligence, today, more than a century after the seminal work of Alfred Binet and the US Army, verbal, nonverbal, and quantitative tests are used to measure intelligence and widely accepted among psychologists and the general public as well and often considered a gold standard for assessment of intelligence (Matarazzo 1992; Stano 2004). As this volume attests, the time has come for a reexamination of the concept

J.A. Naglieri, Ph.D. (✉)
University of Virginia
e-mail: jnaglieri@gmail.com

S. Goldstein
University of Utah, 230 South 500 East, Suite 100, Salt Lake City, UT 84102, USA
e-mail: info@samgoldstein.com

S. Goldstein et al. (eds.), *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts*, DOI 10.1007/978-1-4939-1562-0_30,

of intelligence. The time has come to recognize how early notions and historical precedence have influenced our thinking about intelligence.

Despite the 3,000-year debate about intelligence, traditional IQ as we know it today was formulated just over 100 years ago in 1917 as described by Yoakum and Yerkes (1920) and solidified in 1939 with the publication of the Wechsler-Bellevue scales. The importance of these tests to our society is substantial because they shaped how we define intelligence (Anastasi and Urbine 1997). Results from these tests have influenced the lives of countless children and adults around the world. It is clear that intelligence tests represent one of the most influential contributions made by psychology to society in general. The fact that these tests have had considerable impact on our field is indisputable, as is the fact that they have influenced what we think intelligence *is* and how it should be measured.

The Binet and Wechsler scales have historically measured IQ using tests of similar content. Wechsler's tests were based on Binet's work (Kaufman and Lichtenberger 2000) and the methods used by the US military in the early 1900s (Yoakum and Yerkes 1920). Wechsler included verbal and performance, or nonverbal, tests but not because he intended to measure verbal and nonverbal types of intelligence but rather because it was clear that nonverbal tests helped to "minimize the over-diagnosing of feeble-mindness that was, he believed, caused by intelligence tests that were too verbal in content … [in fact] he viewed verbal and performance tests as equally valid measures of intelligence and criticized the labeling of performance [nonverbal] tests as measures of special abilities" (Boake 2002, p. 396). This view is similar to that described by Yoakum and Yerkes (1920) perhaps best illustrated by the administration procedures for the Army Mental Tests. The general procedure involved assignment of literate draftees to one room, called the Alpha room, and the illiterates to another room, called the Beta room (p. 18). Beta (nonverbal) tests were used because it was known that a person could fail Alpha (verbal and quantitative) tests because of limited educational training and particularly poor skills in reading and writing in English. These persons were then tested with the nonverbal tests to avoid "injustice by reason of relative unfamiliarity with English" (Yoakum and Yerkes 1920, p. 19). Thus, the original intent of verbal and nonverbal tests reflects the practical need that existed 100 years ago and still has relevance today (see Naglieri and Ford 2003) to measure general intelligence.

It is well documented that there is considerable empirical support for the concept of general intelligence as measured by tests such as the Wechsler and Binet (see Jensen 1998 for a review). Perhaps one of the most important sources of validity evidence for traditional IQ tests is the fact that IQ scores are a good predictor of school achievement (Naglieri and Bornstein 2003; Ramsey and Reynolds 2004). But there is circularity in this logic; the verbal tests used to measure intelligence are remarkably similar to the tests used to measure achievement. This problem was true when the tests were first described by Yoakum and Yerkes (1920), and it is still true today for individual as well as group tests of intelligence.

The lack of a clear distinction between ability and achievement has corrupted the very concept of ability in such a way that any child who does not have an adequately enriched educational experience will be at disadvantage when assessed with a so-called "ability" test that involves verbal and quantitative test questions.

The practical importance of separating the content in tests of ability and achievement is particularly salient for children with limited native language skills or those from lower socioeconomic levels where the degree of enrichment in the home is limited. It is well known that poverty or low SES negatively affects students' test performance; high poverty is correlated with low test scores because of issues associated with educational enrichment at home and at school. Thus, many students receive low test scores because of unequal opportunity to learn. Too many of these students – from all racial and cultural backgrounds – are penalized on traditional tests of intelligence and achievement and, subsequently, denied access to gifted education programs and services. Such denial of access is common when tests are highly verbal and highly achievement oriented, as just discussed.

In recent years, however, researchers have begun to examine how effective the general intelligence approach is and indeed to wonder about the limitations of this approach (Naglieri 1999; Sternberg 1988). The verbal/nonverbal approach to conceptualizing intelligence has considerable limitations, especially for culturally and linguistically diverse populations, those with limited English language skills, and children who are experiencing academic problems, like a learning disability. The limited utility of the verbal/nonverbal model for evaluation of specific intellectual problems associated with learning disabled (LD) children's academic failure has led some to argue that intelligence tests are irrelevant to the diagnosis of learning disabilities (Siegle, 1989). Similarly, Kaufman and Lichtenberger (2000) concluded that WISC-III subtest profiles "do not have adequate power on which to base differential diagnosis" (p. 205) for LD or attention-deficit/hyperactivity disorder (ADHD). The recognition that scores on a verbal/nonverbal test of intelligence have not been especially helpful for diagnosis of LD or ADHD (Kavale and Forness 1984; Kaufman and Lichtenberger 2000) should motivate professionals to look elsewhere to understand these conditions. This was one of the goals of this book – to stimulate thought about where we are in the field of intelligence and intelligence testing and what are the most promising suggestions for where we should go from here. We applaud the work of our colleagues in zoology, evolutionary science, psychology, and education to appreciate the genetic and evolutionary roots of intelligence and to move forward to define intelligence. We are confident that the next 50 years of intelligence research will usher a new age in our understanding, evaluating, and enhancing intellectual development.

References

Anastasi, A., & Urbine, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River: Prentice Hall.

Boake, C. (2002). From the Binet-Simon to the Weschsler-Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology, 24*(3), 383–405.

Bracken, B. A., & Naglieri, J. A. (2003). Assessing diverse populations with nonverbal tests of general intelligence. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children* (2nd ed., pp. 243–273). New York: Guilford.

Ford, D. Y. (1996). *Reversing underachievement among gifted black students: Promising practices and programs*. New York: Teachers College Press.

Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist, 52*, 1115–1124.

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *Iowa tests of basic skills*. Itasca: Riverside.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.

Kaufman, A. S., & Lichtenberger, E. O. (2000). *Essentials of WISC-III and SPPSI-R assessment*. New York: Wiley.

Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wechsler scale profiles and recategorizations: Patterns or parodies? *Learning Disability Quarterly, 7*, 136–156.

Lohman, D. F., & Hagen, E. P. (2001). *Cognitive abilities test*. Itasca: Riverside.

Matarazzo, J. D. (1992). Psychological testing in the 21st century. *American Psychologist, 47*, 1007–1018.

Naglieri, J. A. (1997). *Naglieri nonverbal ability test*. San Antonio: The Psychological Corporation.

Naglieri, J. A. (1999). *Essentials of CAS assessment*. Hoboken: Wiley.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21*, 244–260.

Naglieri, J. A., & Ford, D. (2003). Addressing underrepresentation of gifted minority children using the Naglieri nonverbal ability test (NNAT). *Gifted Child Quarterly, 47*, 155–160.

Naglieri, J. A., & Ronning, M. E. (2000). Comparison of white, African-American, Hispanic, and Asian children on the Naglieri nonverbal ability test. *Psychological Assessment, 12*, 328–334.

Naglieri, J. A., Booth, A. L., & Winsler, A. (2004). Comparison of Hispanic children with and without limited english proficiency on the Naglieri nonverbal ability test. *Psychological Assessment, 16*, 81–84.

Pintner, R. (1923). *Intelligence testing*. New York: Henry Holt.

Ramsey, M. C., & Reynolds, C. R. (2004). Relations between intelligence and achievement tests. In G. Goldstein & S. Beers (Eds.), *Comprehensive handbook of psychological assessment* (pp. 25–50). New York: Wiley.

Ramirez, R. R., & Patricia de la Cruz, G. (2002). *The Hispanic population in the United States: March 2002*, Current population reports, P20-545, U.S. Census Bureau, Washington, DC.
Roid, G. (2003). *Stanford-Binet* (5th ed.). Itasca: Riverside.
Stano, J. F. (2004). Test review: Wechsler abbreviated scale of intelligence. *Rehabilitation Counseling Bulletin, 48*, 56–57.
Sternberg, R. (1988). *The triarchic mind: A new theory of intelligence*. New York: Viking Press.
Whitmore, J. R. (1980). *Giftedness, conflict, and underachievement*. Boston: Allyn & Bacon.
Woodcock, R. W., Mc Grew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca: Riverside Publishing.
Yoakum, C. S., & Yerkes, R. M. (1920). *Army mental tests*. New York: Henry Holt.