Innovations in EDUCATIONAL PSYCHOLOGY

Perspectives on Learning, Teaching, and Human Development

Edited by
DAVID D. PREISS
ROBERT J. STERNBERG

SPRINGER PUBLISHING COMPANY

Innovations in Educational Psychology

David D. Preiss, PhD, is an Assistant Professor at the Escuela de Psicología of the Pontificia Universidad Católica de Chile, where he is affiliated with the Measurement Center MIDE UC and the Center for Research on Educational Policy and Practice. He is as well an affiliate member of the PACE Center at Tufts University. Dr. Preiss received his PhD from Yale University, where he attended as a Fulbright Scholar. His main research interests include folk pedagogy, culture and instruction, creativity and writing processes, and the cognitive consequences of technology. He and Dr. Sternberg previously co-edited *Intelligence and Technology: The Impact of Tools on the Nature and Development* (2005).

Robert J. Sternberg, PhD, is currently Dean of the School of Arts and Sciences at Tufts University, where he is also Professor of Psychology. He was previously IBM Professor of Psychology and Education at Yale University. Dr. Sternberg received his PhD from Stanford and is the recipient of 10 honorary doctorates. Moreover, he has won more than two dozen awards for his work. He is a former president of the American Psychological Association and the author of over 1,200 books, articles, and book chapters. His main research interests include intelligence, creativity, wisdom, intellectual styles, and leadership.

Innovations in Educational Psychology

Perspectives on Learning, Teaching, and Human Development

DAVID D. PREISS
ROBERT J. STERNBERG
EDITORS

SPRINGER PUBLISHING COMPANY

New York

Springer Publishing Company, LLC
11 West 42nd Street
New York, NY 10036
www.springerpub.com

Acquisitions Editor: Philip Laughlin
Project Manager: Laura Stewart
Cover Design: Steve Pisano
Composition: Apex CoVantage, LLC

E-book ISBN: 978-0-8261-2163-9

10 11 12 13 / 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Innovations in educational psychology : perspectives on learning, teaching, and human development / [editors] David D. Preiss & Robert J. Sternberg.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-8261-2162-2 (alk. paper)
1. Educational psychology. 2. Critical thinking. 3. Educational innovations. 4. Educational change I. Preiss, David, 1973– II. Sternberg, Robert J.
LB1051.I485 2010
370.15—dc22 2009043751

Printed in the United States of America by the Hamilton Printing Company

Contents

This page intentionally left blank

Contributors

John Baer, PhD
Professor
School of Education
Rider University
Lawrenceville, New Jersey

Ronald A. Beghetto, PhD
Associate Professor
Educational Studies
University of Oregon
Eugene, Oregon

Peter Bryant, PhD
Professor
Department of Experimental Psychology
University of Oxford
United Kingdom

Matia Finn-Stevenson, PhD
Senior Research Scientist, Psychology
Child Study Center
Associate Director
Zigler Center in Child Development and Social Policy
Director
The School of the 21st Century
Yale University
New Haven, Connecticut

Florencia Gomez, MS
Elementary Teacher
Department of Computer Science
School of Engineering
Pontificia Universidad Catolica de Chile
Santiago, Chile

Patricia M. Greenfield, PhD
Distinguished Professor
Department of Psychology at UCLA
Founding Director
FPR-UCLA Center for Culture, Brain, and Development
University of California at Los Angeles

Elena L. Grigorenko, PhD
Department of Psychology
Yale University Child Study Center
Department of Epidemiology and Public Health
Columbia University
Moscow State University
Russia

Patricia Imbarack
Psychologist
Department of Computer Science
School of Engineering
Pontificia Universidad Catolica de Chile
Eduinnova, Santiago, Chile

Patricia T. Kantor, MA
Department of Psychology
Florida State University
Florida Center for Reading Research
Tallahassee, Florida

James C. Kaufman, PhD
Learning Research Institute
Department of Psychology
California State University at San Bernardino

Heidi Keller, PhD
Professor
Developmental and Evolutionary Psychology
University of Osnabrück, Germany

Alex Kozulin, PhD
Director of Research
International Center for the Enhancement of Learning Potential
Jerusalem, Israel

Xiaodong Lin, PhD
Teachers College
Columbia University
New York, New York

Samuel D. Mandelman
Student at Teachers College
Columbia University
New York, New York

Javiera Mena
Psychologist
Pontificia Universidad Catolica de Chile
Santiago, Chile

Sarah Michaels, PhD
Professor
Department of Education
Clark University
Worcester, Massachusetts

María Elena Mora
Sociologist
Department of Computer Science
School of Engineering
Pontificia Universidad Catolica de Chile
Santiago, Chile

Adam J. Naples
Doctoral Student
Department of Psychology
Yale University
New Haven, Connecticut

Terezinha Nuñes, PhD
Professor
Department of Education
Chair in Educational Studies
Harris Manchester College
University of Oxford
United Kingdom

Miguel Nussbaum, MS, PhD
Electrical Engineer
Information and Computer Science
Doktor der Technischen Wissenschaften
Department of Computer Science
School of Engineering
Pontificia Universidad Catolica de Chile
Santiago, Chile

Cathy O'Connor
Faculty Advisor
Language Development
Boston University
Boston, Massachusetts

David R. Olson, PhD
University Professor Emeritus
Ontario Institute for Studies and Education (OISE)
University of Toronto
Toronto, Ontario, Canada

Blanca Quiroz
Assistant Professor
College of Education & Human Development
Texas A&M University
College Station, Texas

Natalia Rakhlin
Postdoctoral Associate
Yale University
New Haven, Connecticut

K. Ann Renninger, PhD
Professor
Department of Educational Studies
Swarthmore College
Swarthmore, Pennsylvania

Lauren B. Resnick, PhD
Learning Policy Center
School of Education
University of Pittsburgh
Pittsburgh, Pennsylvania

Carrie Rothstein-Fisch, PhD
Department of Educational Psychology
California State University
Northridge, California

Alan H. Schoenfeld, PhD
Elizabeth and Edward Conner Professor of Education
Graduate School of Education
University of California at Berkeley

Robert S. Siegler, PhD
Psychology Department
Carnegie Mellon University
Pittsburgh, Pennsylvania

Marcos Singer, MS, PhD
Industrial Engineer
Operations Research
School of Business
Pontificia Universidad Catolica de Chile
Santiago, Chile

Keith E. Stanovich, PhD
Professor
Department of Human Development and Applied Psychology
University of Toronto
Toronto, Ontario, Canada

Paula J. Stanovich, PhD
Professor
Graduate School of Education
Portland State University
Portland, Oregon

Florence R. Sullivan, PhD
Assistant Professor
Program in Educational Technology
School of Education
University of Massachusetts
Amherst, Massachusetts

Alex Torres, MS
Psychologist
Department of Computer Science
School of Engineering
Pontificia Universidad Catolica de Chile
Santiago, Chile

Elise Trumbull, PhD
Formerly Head of the Language and Culture Program at WestEd, San Francisco
Currently Educational Consultant
Oakland, California

Richard K. Wagner, PhD
Department of Psychology
Florida State University and Florida Center for Reading Research
Tallahassee, Florida

Edward Zigler, PhD
Professor Emeritus
Sterling Professor of Psychology
Yale Child Study Center
Yale University School of Medicine
Codirector
Edward Zigler Center Head Start Unit
New Haven, Connecticut

This page intentionally left blank

Preface

More than a decade ago, one of the editors of this book suggested that the field of educational psychology was going through challenging times. He noted that fewer certification programs required courses in educational psychology, relatively few psychologists were involved in initiatives of educational reform, and educational psychology tended to be a marginalized field both in education schools and psychology departments (Sternberg, 1996). Nobody would have expected then that just a decade or so later, issues dealing with education would become so central to several subfields of psychology and related disciplines.

Indeed, over the last decade a number of changes have put educational issues back on the track of basic research. And although the institutional arrangements within many universities may not have necessarily favored this new development, a substantial part of recent educational research has been dealing with what we may consider fundamental scientific questions, such as the nature–nurture debate, the nature of human rationality and intelligence, and the role of culture in human development, among many others. The self-propelled movement of research has triggered some of these changes; others have been triggered by transformations originating outside science.

Thus, new discoveries in genetics and the biological sciences are increasing our understanding of the relation between our innate dispositions and the educational process. Contrary to what we may have expected, these new findings are challenging our understanding of the innate determinants of learning as something static and fixed. On the other hand, many advances in developmental psychology are enriching our understanding of the psychological consequences of schooling; they are also transforming our conception of the learning process and particularly of what we understand by expertise in teaching and learning. In fact, as Sternberg (1996) projected, in recent years the field has experienced an unprecedented growth of studies focused on expert teaching and expert learning (Bransford, Brown, & Cocking, 1999; Torff, 2003;

Torff & Sternberg, 2001). Not surprisingly, some of these initiatives have rapidly translated into classroom recommendations (Sternberg, 2002; Sternberg & Grigorenko, 2003a). Closely connected to these advances are new developments in both individual-differences and cognitive research, which have made important contributions to our understanding of the context-dependent, dynamic, and multifaceted nature of human abilities (Sternberg & Grigorenko, 2003b).

At the same time, findings from educational research have increasingly become a public issue. A central role has been played by the diverse initiatives of educational accountability, which today, for better or for worse, are blossoming around the world. Whereas the No Child Left Behind Act in the United States has sparked most of the recent accountability debate, in other areas of the world, that debate has been encouraged by the results of comparative studies such as the Trends in International Mathematics and Science Study (TIMSS) or the Program for International Student Assessment (PISA), which are informative of the level of achievement reached by different nations in a number of key educational variables (for some illustrations, see Stigler & Hiebert, 1999; Takayama, 2008). Since many of these initiatives comprehensively engage in implicit or explicit forms of teaching assessment, understanding the nature of teaching has become a hot topic in many latitudes and a challenge of utmost importance not only for researchers but also for makers of public policy.

The impact of globalization is making more common the adoption of a comparative stance in educational research (Grigorenko, Jarvin, Niu, & Preiss, 2007). Correspondingly, the last decade has seen renewed interest in the nature of teaching and how it differs across nations and cultures (Alexander, 2000, 2001; Givvin, Hiebert, Jacobs, Hollingsworth, & Gallimore, 2005; Loera, 2006; Preiss, 2009). Part of this interest has been prompted by the development of highly sophisticated technologies that rest on the use of videos for classroom observation, such as those proposed by advocates for TIMSS during the 1990s (Stigler, Gallimore, & Hiebert, 2000) and adopted by many scholars in other parts of the world (e.g., Loera, 2006).

Last but not least, the last decade has seen the proliferation of new technologies whose potential for teaching and instruction are just starting to be fully explored (Mayer, 2003; Preiss & Sternberg, 2005, 2006). More than a decade ago, data provided by the TIMSS reported a substantially low use of computers in an international sample of eighth-grade lessons: Instead of computers, the instructional materials used often in each country were quite “low-tech” and the chalkboard was still prevalent (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). One decade later, the computerization of schools and use of the Internet has advanced significantly; but still, teacher adoption of the new technologies is a cause of concern for many. What is clear is that, far from the Skinnerian idea that teaching can be entirely replaced by universal machines, the new discoveries advanced during the current decade reveal that teachers play an essential role in mediating and making real the intellectual potential of the new technologies (e.g., see Lin, 2001). Still, our understanding of how the traditional and steady routines of teaching and learning can be transformed by using ever-changing technologies is limited.

Given the situation as we have described it, we believed the time was ripe to produce a collection of some of the new findings that are helping to rekindle interest in educational issues among basic researchers and that also are instrumental to communicating these findings among the general public. Particularly, the purpose of this book is to assemble a collection that reviews the current state of the art in several areas of educational research with a view toward contributions that link teaching, learning, and human development and that are, consequently, naturally interdisciplinary. Thus, we expect that this book will help to publicize a number of findings that have high scientific value but do not necessarily receive the attention they deserve. We further hope it will help to advance research in education by providing a single-volume review of important recent discoveries. Finally, we hope that this book will help make education a more attractive endeavor for future scholars by showing how exciting research in this field can be.

Each chapter includes a review of contemporary discoveries the authors deem to be novel and forward-thinking and that are having a major impact on the field. Each author (or group of collaborators) devotes a substantial part of the review to his or her own work. The review includes the nature of the questions that have been addressed with regard to his or her topic, how these questions originated, how they have been commonly answered, how the author and others have addressed these questions, and how the proposed approach compares to and contrasts with others. Authors have been encouraged to place their reflections in an interdisciplinary context and to relate teaching, learning, and human development. The chapters have been written in a way that is understandable to first-year graduate students in psychology or education and related disciplines. Yet we are sure that the book will appeal to a diverse audience.

The book is organized into five sections. We briefly describe each in turn.

Section I is devoted to new research on individual differences in human development. The first chapter, by Elena L. Grigorenko, Samuel D. Mandelman, Adam J. Naples, and Natalia Rakhlin, discusses what the authors define as "the seemingly eternal but incorrectly posed dichotomous question of genes versus environment" from the point of view of its educational implications. As this chapter illustrates, we may need to go beyond the mere decomposition of variance if we are to understand the dynamic nature of the interactions between genes and environment and what education has to do with it. The second chapter in this section, by Richard K. Wagner and Patricia T. Kantor, is devoted to reviewing recent findings regarding the causes of dyslexia, its identification, and ways to help those individuals affected by it. For many years, dyslexia has drawn attention of researchers for both basic and practical reasons. On the one hand, problems in reading are informative of many basic cognitive processes. On the other hand, as Wagner and Kantor note, the vast majority of students who receive assistance for specific learning disabilities have problems in reading, and problems in reading have long-term effects in human development. So a chapter on dyslexia could not be omitted in a book devoted to the dissemination of new research findings in education.

The second section is devoted to chapters that deal with the complex relation between schooling as a process of deliberate instruction and human development as a process of mental growth, broadly understood.

The first chapter in this section, by Terezinha Nuñes and Peter Bryant, discusses several insights from everyday knowledge for mathematics education. The authors note that people with little school instruction typically solve proportion problems in everyday life by setting values in one-to-many correspondence across variables. The authors investigated the origin and uses of this schema of reasoning by schoolchildren. This schema often goes unnoticed in school, where multiplication is taught as repeated addition. The chapter therefore discusses the difference between these two conceptions of multiplication. Research shows why one-to-many correspondence reasoning offers a good foundation for teaching children about proportions in school.

The second chapter of this section, by Xiaodong Lin, Robert S. Siegler, and Florence R. Sullivan, describes several studies showing that not all students view learning as their primary goal in school and that students' goals vary between and within cultures. The authors show that optimizing students' learning requires attention to learner's goals and values and that instruction leading students to adopt the goal of deeply understanding the material that is being taught can produce superior learning.

The third chapter of this section, by K. Ann Renninger, focuses on the critical role of interest in learning and development. Renninger shows that interest will develop when the learning environment seeks to cultivate learner interest in relation to phase of interest, motivational profile, and principled knowledge of content. Learners need to feel that their efforts are respected and to receive ongoing support to recognize, take advantage of, and seek opportunities to think and reengage.

The third section is devoted to studies on what we call the refinement of mind. In the first chapter of this section, James C. Kaufman, Ronald A. Beghetto, and John Baer present a new model of creativity and discuss the phenomenon of creative polymathy. They consider multicreative potential in light of their "Four C" model and offer an alternate possibility to general and domain specific models of creativity by drawing on the Amusement Park Theoretical (APT) model.

The second chapter of this section, written by Lauren Resnick, Sarah Michaels, and Cathy O'Connor, discusses the process of reasoning within classrooms and how children can be taught to give reasons for their answers. The authors introduce methods of creating these competencies of what they call "accountable talk," which have been tested by the authors within face-to-face, teacher-led learning environments. The authors present results showing that students who learn in classrooms guided by accountable talk standards are socialized into communities of practice in which respectful and grounded discussion prevails.

The third chapter of this section, by Keith E. Stanovich and Paula J. Stanovich, presents a tripartite model of mind that explains why rationality is a more encompassing construct than intelligence. Similarly, the authors subsume the construct of critical thinking under the construct of rationality as well. According to the authors, creating a generic model of the mind that has rationality as an overarching construct, which integrates critical thinking and intelligence, has the considerable benefit of placing the construct of critical thinking within contemporary cognitive science.

The fourth section of the book is devoted to the experiences of teaching and learning and their cultural grounds. The opening chapter of the section, by Alan H. Schoenfeld, describes a model of teaching-in-context, which explains, on a line-by-line basis, decision making by teachers during hour-long classroom lessons. The existence of such models provides tools for examining and improving teaching as well as the possibility that decision making in other professions can be comparably modeled.

The second chapter of the section, by Carrie Rothstein-Fisch, Patricia M. Greenfield, Elise Trumbull, Heidi Keller, and Blanca Quiroz, describes the relevance of the value systems of individualism and collectivism for learning, development, and education. The authors focus their attention on situations where the home culture of collectivistic children opposes the individualistic culture of schools, which creates a need for educational intervention. One such situation is that of Latino children in the United States. An educational intervention addressing their needs, the Bridging Cultures Project, is presented.

The third chapter of the section, by Miguel Nussbaum, Florencia Gomez, Javiera Mena, Patricia Imbarack, Alex Torres, Marcos Singer, and María Elena Mora, discusses ways of transforming the learning process into an interactive and collaborative process by using handheld devices with wireless networks. The authors discuss a methodology for technology-supported small-group collaborative learning that uses assessment to adapt teaching to meet learning needs. They report the results of a detailed qualitative and quantitative study of the methodology, which demonstrates that communication and social abilities are attained with an improvement in learning.

The last chapter of the section, by David Preiss, presents a tentative operationalization of the construct of folk pedagogy, discusses its application to a context other than the one where this construct originated, and presents three empirical illustrations based on studies done in Chile and inspired by this theoretical framework. Preiss's results provide clear support for the hypothesis that there are country-specific patterns of teaching. Preiss concludes that future public policy initiatives for educational reform should take into consideration the cultural basis of the teaching patterns they want to transform and how permeable these patterns are to initiatives originating outside of classrooms.

The last section addresses the old theme of school reform by introducing three new initiatives in this area. The first chapter of the section, by Edward Zigler and Matia Finn-Stevenson, describes the School of the 21st Century (21C), a preschool program that provides not only care and education for preschoolers but also other family support programs beginning at the conception of the child and throughout the school years. This program has been implemented in more than 1,300 schools, enabling educators to address the needs of children and families. In this chapter, the authors describe 21C and review its research base, implementation, and impact, showing how the program is transforming education in the United States.

The second chapter of the section, by Alex Kozulin, focuses on four areas of research with a great potential of use for initiatives of school reform. These are the general conceptualization of the goals and framework of the educational process, the multicultural classroom, the evaluation of students' learning potential, and emerging research on differences between such fundamental processes as thinking and learning.

The next chapter of the section, by Robert J. Sternberg, advances the WICS model as a possible common basis for the development of skills and attitudes in college. WICS is an acronym standing for wisdom, intelligence, and creativity, synthesized. Wisdom, intelligence, and creativity are sine qua nons for the citizens and professionals of the future and really for anyone who wishes to achieve meaningful success in his or her life. Sternberg discusses each of these attributes and describes methods for developing and measuring the attributes.

In the closing chapter, on the subject of adopting a perspective based on the development of education as a discipline, David R. Olson integrates all the findings presented in this book. He examines the possible ramifications of this work for educational reform, educational practice, and educational research.

David Preiss's work on the preparation of this book was supported by the Center for Research on Educational Policy and Practice, Grant CIE01-CONICYT and Grant Number 11060389 from FONDECYT (Fordo National de Desarrollo Cieatifiro & Technologica). He dedicates this book to his daughters Ilana and Meital, whom he expects may benefit at their schools, present and future, from practices inspired by good educational research. Bob Sternberg dedicates the book to his graduate advisor, Gordon Bower, who has inspired him throughout his career to apply his ideas from psychology to education.

David D. Preiss
Robert J. Sternberg

REFERENCES

Alexander, R. (2000). *Culture and Pedagogy. International comparisons in primary education.* London: Blackwell Publishing.

Alexander, R. (2001). Border crossings: towards a comparative pedagogy. *Comparative Education,* 507–523.

Bransford, J., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Givvin, K. B., Hiebert, J., Jacobs, J. K., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review, 49*(3), 311–343.

Grigorenko, E. L., Jarvin, L., Niu, W., & Preiss, D. (2007). Is there a standard for standardized testing? In P. C. Kyllonen, R. D. Roberts & L. Stankov (Eds.), *Extending intelligence: Enhancement and new constructs* (pp. 157–182). Mahwah, NJ: Erlbaum.

Lin, X. (2001). Reflective adaptation of a technology artifact: A case study of classroom change. *Cognition and Instruction, 19*(4), 395–440.

Loera, A. (2006). *La práctica pedagógica videograbada. [Teaching practices videotaped].* México, DF: UPN.

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13,* 125–139.

Preiss, D. D. (2009). The Chilean instructional pattern for the teaching of language: a video-survey study based on a national program for the assessment of teaching. *Learning and Individual Differences, 19,* 1–11.

Preiss, D. D., & Sternberg, R. J. (2005). Technologies for working intelligence. In *Intelligence and technology: The impact of tools on the nature and development of human abilities* (pp. 183–208). Mahwah, NJ: Erlbaum.

Preiss, D. D., & Sternberg, R. J. (2006). Effects of technology on verbal and visual-spatial abilities. *International Journal of Cognitive Technology, 11*(1), 14–22.

Sternberg, R. J. (1996). Educational psychology has fallen, but it can get up. *Educational Psychology Review, 8*(2), 175–185.

Sternberg, R. J. (2002). Raising the achievement of all students: Teaching for successful intelligence. *Educational Psychology Review, 14*(4), 383–393.

Sternberg, R. J., & Grigorenko, E. L. (2003a). Teaching for successful intelligence: Principles, procedures, and practices. *Journal for the Education of the Gifted, 27*(2–3), 207–228.

Sternberg, R. J., & Grigorenko, E. L. (2003b). *The psychology of abilities, competencies, and expertise.* New York: Cambridge University Press.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist, 35,* 87–100.

Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States.* Washington, DC: U.S. Government Printing Office.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap. Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.

Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education, 44*(4), 387–407.

Torff, B. (2003). Developmental changes in teachers' use of higher order thinking and content knowledge. *Journal of Educational Psychology, 95*(3), 563–569.

Torff, B., & Sternberg, R. J. (2001). *Understanding and teaching the intuitive mind: Student and teacher learning.* Mahwah, NJ: Erlbaum.

PART I

Individual Differences in Human Development

This page intentionally left blank

1

Juxtaposing Psychological, Educational, and Genomic Sciences: An Emerging Platform for Interpreting Individual Differences in the Classroom

ELENA L. GRIGORENKO, SAMUEL D. MANDELMAN, ADAM J. NAPLES, AND NATALIA RAKHLIN

In science, ideas come and go. Some ideas are driven out by an accumulation of opposing scientific evidence; others seem to take forever to get rid of, no matter what the evidence suggests. The nature–nurture controversy—the long-standing debate over whether heredity (i.e., influences from within, inherited from or transmitted by previous generations) or environment (i.e., influences from the outside, encountered throughout the life span) is more important in human development—is just one of these hangers-on. In the field of scientific inquiry, no matter how much evidence accumulates on the etiology of human behavior, this controversy continues to be invalid as a logical premise. In other words, the nature–nurture debate continues to exist, despite the false dichotomy it poses.

A clearer understanding of the relationships between heritable and nonheritable factors and behavior is something that psychologists and educators have been interested in for a long time. Since the early work of Sir Francis Galton (Galton, 1869), the question of what is hereditary and what is not has been asked about many human traits in many contexts. Approximately a century and a half later, it appears that, with the exception of inconsequential or poorly measured human traits, pretty much all traits demonstrate some degree of heritability. The interpretation of heritability as a statistic is typically focused on two of its aspects:

(a) whether it is statistically different from zero and (b) what its magnitude is when its confidence interval is taken into account. When the heritability is statistically different from zero, the statistic indicates that the population/sample–phenotypic variability is attributable in part to genetic population/sample variability. The magnitude of the heritability estimate itself indicates the magnitude of such attribution.

A few decades of research in behavior genetics, an interdisciplinary field that attracts both psychologists and geneticists, indicate that the majority of complex human traits, especially those associated with cognitive functioning and educational achievement, are heritable (meaning that at least a portion of the phenotypic variance in the trait is associated with genetic variance and that that portion is statistically different from zero). However, similarly, the majority of such traits are also susceptible to environmental impacts (i.e., the phenotypic variance is also, at least in part, attributable to environmental variance). In fact, it is difficult to find a trait—at least one that has been approached with well-designed studies and reasonably large samples—for which the variance underlying the observed individual differences is primarily either genetic or environmental. Typically, there is a mixture of influences, and estimates of the ratio for this mixture can fluctuate developmentally, but rarely if ever is it either/or. Thus a substantial investment of time and money spent in attempting to resolve the controversy has resulted in the realization that it cannot be resolved. So, why does the nature–nurture issue keep popping up?

It has been suggested that the long life of this controversy is attributable to the truly complex reality of the relationships among the genome, the brain, and behavior; this reality has been described as unfolding on a number of time scales, ranging from a million years of human evolution to split-second decision making in extreme situations (Robinson, 2004). The multidimensional connection between the genome and behavior is indirect and mediated by the brain, which, in turn, is built by proteins encoded by the genes, the essential building blocks of the genome. However, the brain's "construction zone" is not situated in a galactic vacuum but rather localized within the context of a particular environment that, in turn, impacts the machinery of the genome and the brain's edifice and function (Robinson, Fernald, & Clayton, 2008). This environment changes over time, calling for matching changes in the genome by both impacting it directly and signaling it through the brain. Thus what should replace the concept of the nature–nurture controversy is the concept referred to by Robinson as "the dynamic genome" (Robinson, 2004, p. 399)—an open structure whose function is shaped by both heritable (i.e., "old," preexisting in some form, transmitted across generations) and nonheritable (i.e., imposed anew, emerging in a particular individual) forces.

What does this new perspective mean for the field? What will happen to the seemingly eternal but incorrectly posed dichotomous question of genes versus environment? Here, it is argued that this question can never be answered and thus should nevermore be asked. When it was asked by Galton in the middle of the 19th century, the value of this question was very different from its value today, here at the beginning of the 21st century. Asked then, it seemed priceless, opening a new direction of scientific inquiry; asked now, it is meaningless.

Galileo Galilei's action in dropping balls from the Tower of Pisa and Isaac Newton's observation of falling apples at Trinity College, Oxford, stimulated the formulation of the laws of gravity. Now, if a child asks why apples fall down instead of up or float in the air, a clear answer can be provided based on the modern understanding of the laws of gravity. There is no need for humanity to deal with balls and apples for the sake of understanding gravity: we have captured its meaning! Similarly, we now also know the answer to the genes-versus-environment question and thus should put the nature–nurture controversy to rest. Yet "should" is not always easily converted into "is."

The statement that cutting-edge science has moved far beyond the nature–nurture controversy, of course, does not mean that there are no other interesting questions to ask in our quest to understand *how* the dynamic genome realizes itself in human behavior. Similarly, although the laws of gravity have been formulated in principle, there are endless subsequent important qualifiers and addenda to these laws. Searching for answers to the *how* question is, arguably, the most productive way of generating new data on the connection between the genome and behavior. This broad question can be reformulated in many different ways. Among these rephrased questions are these: *How did evolution make our genomes the way they are*? *How does an individual genome form from the genomes of its progenitors*? *How does it change through stages of life, distancing itself from its original genome*? And there are many others. This broad question of *how* (or one of its variants) is central to the discussion in this chapter, the main purpose of which is to draw four observations from the data that are accumulating in response to the *how* query. These observations are made with references to both human and animal data, and follow the footsteps of the field in an attempt to understand the connection between the dynamic genome and human development in general and cognitive development in particular. Of particular interest here is the formulation of hypotheses with regard to the contribution of education to this connection.

A pronounced limitation of this overview is its inability to cite literature that is directly related to education; there is virtually no such literature. Yet there are massive amounts of related literatures on humans and animals that form the foundation for formulating hypotheses with regard to the connections between these literatures and education. Thus, this chapter is structured around the discussion of four observations and concludes by linking these observations to issues of education. These observations are on (a) the role of the dynamic genome in the evolution of uniquely human traits; (b) associations between the genome, the brain, and subsequently cognition; (c) attempts to localize cognition in the genome; and (d) the correspondences between the degree of complexity of particular behaviors and the genetic mechanisms governing them.

Of note is that, in the 20th century, psychology and education, as scientific disciplines, have been predominantly driven by research paradigms that rest on the assumed necessity of formulating a priori hypotheses before collecting data and capitalizing on inferential–statistical frameworks of data analyses. Quite to the contrary, the genomic sciences, especially within the last decade, prior to and since the sequencing of the human genome, have been predominantly inspired by exploratory paradigms and are open to data mining and the engagement of posterior probabilities approaches. The illustrations of the connections between the dynamic genome and complex behaviors, presented below, highlight the impact of explorations characteristic of modern genomic sciences. They do not necessarily fit or verify higher-order theories; they try to capture and describe characteristics of the dynamic genome with regard to its relationship to complex human behaviors.

OBSERVATION 1: EVOLUTION AND UNIQUELY HUMAN COMPLEX TRAITS

For many subfields of psychology (e.g., clinical and social in particular), it is important to know how certain human behaviors evolve and what the “wild” (i.e., the initial or ancestral) versus acquired (i.e., modified by human civilization and culture) behaviors are. This evolutionary metric can aid both in understanding typical and atypical behaviors and in defining the parameters of the modifiability of human behavior in response to social pressures. At what behavioral junctions did humans acquire the needed degrees of freedom to deviate from the repertoire of animal behavior? Which genes provided us with the opportunity to gain distance from our "wild" animal genome and attain such traits as free will and creativity, which are as nonanimal as any human traits that we know of? And how did they gain this distance?

A tremendous degree of overlap between the structures of the genome of high apes and modern *Homo sapiens* suggests that evolution is rather unlikely to operate in straightforward and direct ways, building new genetic structures to support new human-specific functions. Given the compact size of the human genome, one hypothesis is that a gain of function is related not to an increase in the size of the genome itself through the addition of new genetic material but rather to reusing existing genetic material in novel ways. This hypothesis assumes that there are some regulatory elements in the genome that guide the usage of the genetic materials, allowing for its differential transcription (i.e., differential gene expression).

To illustrate, a recent investigation (Prabhakar et al., 2008) focused on a particular type of such DNA regulatory material, the so-called cis-regulatory elements, a DNA region that regulates the transcription of genes located on the same DNA strand. More specifically, using bioinformatic approaches, these researchers identified a particular cis-element they described as "the most rapidly evolving human noncoding element yet identified" (p. 1346). Specifically, they identified a 546-base-pair piece of DNA termed human-accelerated conserved noncoding sequence 1 (*HACNS1*). This sequence is conserved and nonvariable in all sequenced vertebrate genomes; however, during the period of approximately 6 million years since the evolutionary differentiation between the human and the chimpanzee, it has accumulated 16 human-specific changes in the DNA structure. A rapid divergence such as that exhibited by *HACNS1* is highly unlikely unless a positive selection pressure that might have altered the function of *HACNS1* is assumed. Researchers (Prabhakar et al., 2008) have presented evidence that *HACNS1* appears to be acting as an enhancer of gene expression that is localized to limbs. Their argument is especially appealing because of their molecular comparative analyses, according to which chimpanzee and rhesus macaque copies of this element do not result in such a comparative enhancement. In order words, the pattern of gene expression enhanced by this element might be directly related to the limb/thumb patterning that differentiates humans from their closest evolutionary relatives.

Based on this observation, it is possible to assume that similar mechanisms might have been involved in the evolution of language. Although currently there are no clear molecular "candidates" for the emergence of human language or its precursors, there are threads of indirect evidence supporting the general idea of the involvement of the genome in such differentiation. Two relevant ideas are discussed here.

First, it has been hypothesized that the differentiation of human languages corresponds to the differentiation of genetic variation between different human populations (Cavalli-Sforza, Piazza, Menozzi, & Mountain, 1988). Researchers have attempted to verify this initial hypothesis, but with mixed success (e.g., Barbujani & Pilastro, 1993; Barbujani & Sokal, 1990; Dupanloup de Ceuninck, Schneider, Langaney, & Excoffier, 2000; Excoffier, Smouse, & Quattro, 1992; Poloni et al., 1997; Rosser et al., 2000). The range of opinions and data collected in an attempt to verify this hypothesis has been broad, extending from the view that "genetic and linguistic features are tightly correlated" (Cavalli-Sforza et al., 1988) through the idea that "these features are correlated, but only in certain regions of the Old World" (Nettle & Harriss, 2003), to "these features are not correlated" (McMahon, 2004). According to these hypotheses, genetic and linguistic variations in different populations are correlated but might not necessarily be causally related.

Second, it has been hypothesized that there might be causal relationships connecting the genetic variation in a given population and certain linguistic feature(s) that characterize the language used by this population. For example, it has been proposed that the emerging variation in certain genes—such as genes involved in such functions as corticogenesis (e.g., *H. sapiens* abnormal spindles mRNA, *ASPM* or *MCPH5,* located at 1q31, and *Microcephalin, MCPH1,* located at 8p23, genes)—might be causally related to linguistic variation, but in an indirect fashion. Specifically, it is possible that the variation in these genes is related to differences in the size and organization of the cerebral cortex, which, in turn, might be related to a subtle cognitive bias in the processing and acquisition of linguistic tone and, correspondingly, to the presence or absence of tone in world languages (Dediu & Ladd, 2007). This hypothesis appears also to be supported by evidence pointing to the young age and relatively quick propagation of the derived variants of both *ASPM* and *Microcephalin* (5,800 and 37,000 years respectively) and suggests, as in the example of *HACNS1* above, the presence of a positive selection force that capitalized on a particular genetic feature magnifying the possibility of certain human traits to arise (Evans et al., 2005; Mekel-Bobrov et al., 2005).

Although the lines of investigation described above are of great interest, they face the major challenge of being translated from generic mechanisms to illustrations pertaining to specific human traits. And this translation is at its very beginning and not without controversies. For example, researchers (Mekel-Bobrov et al., 2007) have performed a large-scale genetic association study attempting to investigate the connection between the adaptive alleles of the *MCPH1* and *MCPH5* genes and normal variation in several measures of IQ in a sample of approximately 3,000 participants. They did not reveal an association that could have supported the connection between the recent adaptive evolution of either the genes or changes in IQ. Yet another study attests to the presence of genetic associations connecting variation in the *MCPH1* gene with both linguistic and nonverbal intelligence (Christiansen, Kelsey, & Tomblin, 2008).

Thus the findings are perhaps contradictory. Yet they are very exciting in terms of their potential to help identify both coding and noncoding regions in the human genome, whose evolution might underlie the emergence of uniquely human traits (e.g., language, as discussed above). In short, the findings presented here permit the observation that the emergence of human features (e.g., the differentiation of the prehensile thumb and the emergence of human language) are related to specific changes in the genome. It appears that different regions of the human genome differentiate at different rates, both presenting the link between *H. sapiens* and the evolutionary tree and allowing the species to distance itself from this tree by acquiring uniquely human features. It is possible that it is those regions in the human genome that accumulate rapid structural changes that are particularly pertinent to understanding the connection between the genome and education. Having identified such rapidly evolving regions, the field might be able to focus on them in a more systematic way in its attempt to understand the genetic bases for individual differences in such traits as the acquisition of literacy and numeracy.

OBSERVATION 2: BRAIN STRUCTURE, GENES, AND COGNITION

One of the axioms of modern cognitive science is that variation in cognitive performance is related to variation in the brain. Generally speaking, the variation in the brain is attributable to differences in (a) brain size, (b) brain structure, and (c) brain activation patterns. Correspondingly, researchers attempt to correlate individual differences in cognition and cognitive tasks with all these sources of variation (for a more detailed discussion of this issue, see Mandelman & Grigorenko, in press).

There is a substantial literature that connects cognition with brain size and structure. For example, the average correlation between brain size and intelligence has been reported to be approximately .33 (McDaniel, 2005). It also has been argued that this correlation is attributable to genetic factors (Posthuma et al., 2002; Posthuma, de Geus, & Boomsma, 2003; Toga & Thompson, 2005; Winterer & Goldman, 2003). This hypothesis was suggested in particular based on the observation that the correlation between the brain properties of monozygotic twins and their intelligence is higher than in dizygotic twins (Posthuma et al., 2002). Summative interpretations of the literatures on intelligence and the brain (e.g., Hulshoff Pol et al., 2004) point to the connection between IQ and the volume and density of gray and white matter in the brain network that engages the regions of the right medial frontal, occipital, and right parahippocampal (gray matter) regions of the brain and the regions of the superior occipitofrontal fascicle and corpus callosum (white matter connecting the corresponding gray matter regions) of the brain.

Developmentally, increases in gray and white matter volume and density (i.e., increases in cortical thickness) are associated with brain (and correspondingly cognitive) maturation. This maturation is the result of numerous morphological changes, including the formation of new neuronal connections by dendritic spine growth as well as changes in the strength of existing connections (Chklovskii, Mel, & Svoboda, 2004), and axonal remodeling and increased soma and nuclei of neurons (Kleim, Lussnig, Schwarz, Comery, & Greenough, 1996). These changes have been attributed to both genetic and environmental effects that unfold in a complex systematic fashion (Shaw, 2007). Although the causal hypothesis connecting brain maturation to the development of intelligence has been rooted in animal literature and supported primarily by it, there are many correlational studies in humans that indirectly buttress this hypothesis.

Specifically, postmortem studies indicate that the brains of individuals with higher IQ and higher levels of education are characterized by a greater number of dendrites and more dendritic branching (Jacobs, Schall, & Scheibel, 1993; Jacobs & Scheibel, 1993) compared with individuals with very low IQs (Huttenlocher, 1991). Yet recent evolutionary analyses of the covariation between brain size and intelligence indicate that there is an evolutionary preference for strong stabilizing (average is better) selection (Miller & Penke, 2007). Thus, although within a given population there is a tendency for intelligence and brain size to correlate, there is no evidence that evolution systematically "promotes" big brains and/or high levels of intelligence. One possible hypothesis here might be that this tendency toward being average in terms of having more stable, biologically controlled traits (e.g., brain size and structure) might explain the greater flexibility and diversity in more dynamic traits such as styles of information processing.

Similarly, there is ongoing interest in behavior–genetic studies of brain anatomy. Although this is a relatively new line of research, the data accumulated so far have indicated that heritability estimates vary for different regions and different ages (Lenroot et al., 2009). They also appear to vary depending on how fine-grained the investigated regions are, with heritabilities being smaller (.00–.50) for smaller regions, such as the thalamus and hippocampus (Wright, Sham, Murray, Weinberger, & Bullmore, 2002), and higher for large brain subdivisions (.60–.80) such as the frontal, parietal, and temporal lobes (Baare et al., 2001; Geschwind, Miller, DeCarli, & Carmelli, 2002; Wallace et al., 2006; Wright et al., 2002). There have also been attempts to investigate the structure of genetic variance by means of factorial techniques, but the results have been inconsistent, with some suggesting (Lenroot et al., 2009) and others negating (Wright et al., 2002) the presence of a major single factor accounting for the majority of genetic variability in the brain anatomy.

Similarly, studies of patterns of brain activation in people who are engaged in cognitive tasks have also produced a pattern of results that is rather difficult to interpret. It has been reported that individual differences in patterns of brain activation in people engaged in working-memory (N-back) tasks are attributable, at least partially, to genetic variation (Blokland et al., 2008). Likewise, it appears that heritability estimates of lateralization for such functions as language (Sakai, Miura, Narafu, & Muraishi, 2004; Sommer, Ramsey, Mandl, & Kahn, 2002), although statistically significant, differ depending on a number of "other" variables, such as handedness (Sommer et al., 2002) and training relevant to developing the tested cognitive function (Sakai et al., 2004). Yet, the activation patterns of the ventral visual cortex in response to language stimuli (pseudowords) were reported as not heritable, whereas the neural activity outlines in the same area, but in response to face and place stimuli, showed genetic influences (Polk, Park, Smith, & Park, 2007). Researchers (Matthews et al., 2007) also reported the presence of genetic influences on the patterns of activation of the anterior cingulate during interference processing.

Thus, in general, the findings in this domain of research are once again rather contradictory. A possible new avenue of inquiry has recently been introduced by an elegant study that blended an investigation of individual differences in activation patterns in response to a particular task, regardless of what that task was, with an objective of differentiating brain networks that are engaged in response to different tasks, in this case, frontoparietal spatial networks and language-related networks (Koten et al., 2009). Perhaps surprisingly, the highest genetic influences were seen not in conjunction with the details (i.e., intensity or temporal resolutions) of patterns of activation but in the general selected cognitive strategy. Interpreting their data, researchers suggested that genetic effects are more related to qualitative, strategy, and style-based differences than to quantitative differences in patterns of activation.

This finding rings true for educators. The educational literature is replete with observations about individual styles of learning that students demonstrate in the classroom. It is these holistic stylistic preferences that differentiate students, not specific peculiarities like the details relating to how they acquire the alphabet or hand-write the letter "o." Learning about how the genome contributes to the formation of such stylistic preferences might enhance the field's understanding of how to address these styles in a classroom. It might also aid in focusing pedagogical efforts on those traits that are more modifiable and more open to intervention than others.

OBSERVATION 3: LOCATIONS IN THE GENOME FROM WHICH COGNITION ORIGINATES

For the last two decades or so, researchers have been engaged in a search for the specific genes involved in the etiology of intelligence and intellectual abilities and disabilities. Such searches usually unfold in one of two ways: as exploratory whole-genome investigations/screens (often also referred to as "scans") or as hypothesis-driven studies of candidate regions in the genome or candidate genes (for a more detailed discussion of this issue, see Mandelman & Grigorenko, in press).

Up to now there have been six genomewide scans for genes contributing to intelligence and cognition (Butcher, Davis, Craig, & Plomin, 2008; Buyske et al., 2006; Dick et al., 2006; Luciano et al., 2006; Posthuma et al., 2005; Wainwright et al., 2006). The results of these scans are quite variable but there are interesting partial overlaps. Specifically, the findings coincide in regions on chromosomes 2q (for 4 out 6 studies), 6p (for 5 out of 6 studies), and 14q (for 3 out of 6 studies). These overlapping regions have been putatively interpreted as indicative of the presence of genes that could explain some of the variance in IQ.

A number of observations can be derived from these results. The first pertains to the variety of the measures used in these studies. In fact, only one study (Butcher et al., 2008) utilizes an indicator that was referred to as measuring the general factor of intelligence, the *g*-factor. The remaining studies used a range of indicators of both achievement and abilities and generated a wide spectrum of findings, allegedly implicating 13 (out of 22) autosomal chromosomes, 5 of which reportedly demonstrated the signals on both arms, *p* and *q*. Thus, between all of these phenotypes and all of these regions, the resulting picture is rather difficult to interpret. Second, of note is the observation that the magnitudes of the presented statistics and *p* values are rather modest. When such effect sizes are estimated (e.g., as in Butcher et al., 2008), they are reported to be very low (topping out at 0.4%). Third, it is important to note that these studies are not independent of each other. They are collectively presented by four groups (two of which, the Dutch and the Australian group, have also published on samples together [Posthuma et al., 2005]), and it appears that there is a substantial overlap in the samples of participants (e.g., Buyske et al., 2006; Dick et al., 2006; Luciano et al., 2006; Posthuma et al., 2005; Wainwright et al., 2006). Given that the presentations are split based on the availability of a complete (or, in some cases, semicomplete) IQ battery versus the availability of specific subtests from IQ tests and/or other cognitive tests and different inclusion/exclusion criteria (e.g., as in Buyske et al., 2006; Dick et al., 2006, Luciano et al., 2006; Wainwright et al., 2006), one may ask whether any of the reported signals would survive if a conservative but traditional approach to correcting for multiple comparisons were applied. Fourth, it is important to note that these studies used a variety of designs and methodologies, analyzing both pooled DNAs for groups of individuals (i.e., mixing DNA from different individuals in one tube) and individual DNAs (i.e., keeping DNA from different individuals in different tubes), recruiting family members and singletons, and covering the genome with genetic markers at highly variable densities. All of these "differences and similarities" must be carefully taken into account in considering the patterns of consistencies and inconsistencies in these findings. Fifth, none of these studies were specifically created to investigate the genetic bases of intelligence, however defined. In fact, the same genetic data were used to investigate linkage/association with multiple other phenotypes in different subsamples of the same samples. At this point, the impact of such reutilization of data on inferential statistics has not been carefully appraised, but there have been concerns in the literature regarding the effect of such reutilization on *p* values, the definition of replicability, and the generalizability of the results (e.g., McCarthy et al., 2008).

In summary, although these scans present interesting data, the reported findings must be interpreted with caution. Yet these studies are considered interesting enough to suggest that further investigations of the genetic bases of intelligence (broadly defined) are warranted.

Although such genomewide scans have not generated specific candidate genes for intelligence, other types of studies have implicated specific genetic regions or specific genes. A comprehensive review of so-called candidate genes for cognition, of which there are many, is beyond the scope of this chapter (for a more detailed discussion, see Mandelman and Grigorenko, in press); thus, only main conclusions relevant to the main argument of this chapter are outlined. First, of note is the fact that these candidate genes are numerous, with sparks of evidence both supporting and disputing the involvement of pretty much each of these genes with the etiology of cognition. Second, these genes are of rather diverse functions (i.e., belonging to families of genes connected to different groups of proteins with known functions in the brain), indicating, presumably, that the genetic pathways to cognition are complex or multilayered. Third, in many of these studies of genes and cognition, the behavioral variables of interest are defined beyond IQ. In fact, they encompass a whole gamut of characteristics of intelligence (verbal and nonverbal, at the minimum) and cognition (e.g., executive functioning, creativity, working memory, and IQ itself). Finally, the participants in whose samples these candidate genes have been investigated differ in age, suggesting that these diverse findings might reflect some developmental variability that is yet to be understood.

Nevertheless, although there have been no or only failed attempts to replicate the findings from some of these candidate-gene studies, there is a degree of consistency in correlating variation in selected genes with variation in cognition. The establishment of these specific associations between genes and cognition is a fundamental breakthrough, a switch from the hypothetical decomposition of variance that was characteristic of earlier heritability studies to a firm “grounding” of these heritabilities in the genome. The hope is that, by understanding the functions of these genes and their interactive proteins networks, the field will gain some additional understanding of how the general biological (and the specific genetic) machinery of intelligence, cognition, learning, and academic skills works. Such discoveries might lead to discussions of the degree to which human potential can be maximized, but with pedagogical and perhaps pharmacological efforts.

OBSERVATION 4: THE COMPLEXITY OF THE STRUCTURE OF COGNITIVE PROCESSES DOES NOT NECESSARILY CORRELATE WITH THE COMPLEXITY OF THE CORRESPONDING GENETIC MECHANISMS

Quite often, in formulating initial hypotheses, researchers start by generating default assumptions that are maximally parsimonious. For example, one such assumption is that a simple behavior or a cognitive process is more likely to be guided by a simple genetic mechanism and a more complex behavior or process is more likely to be controlled by a more complex genetic mechanism. Yet although it is logical, this assumption does not seem to be consistently supported by the literature on how the dynamic genome behaves.

On the contrary, the literature today (for a review, see Robinson, 2004) contains rather surprising examples of complex sustainable behaviors in animals (e.g., pair bonding, foraging, and care of offspring), which are apparently controlled by relatively simple genetic mechanisms. Of interest is the fact that some of these behaviors involve molecules known to be present in humans and therefore likely to be involved in similar types of human behaviors. This literature delivers many interesting observations, two of which appear to be particularly remarkable. The first is that really complex behaviors can be controlled by a single polymorphism in a single gene. The second is that relatively simple (or low-level, componential) behavior can require the coordinated action of many genes.

Consider relevant illustrations. Researchers have explored the molecular bases of mating preferences in two species of voles, the monogamous prairie vole and the polygamous meadow and mountain vole (for a review, see Donaldson & Young, 2008). It has been reported that monogamous versus polygamous behaviors can putatively be associated with a highly polymorphic, complex, repeat-containing DNA element (referred to as a microsatellite) located in the so-called 51-region of the vasopressin receptor, or *AVPR1A,* gene (Donaldson & Young, 2008). Moreover, there is growing evidence in the literature that genetic variation in the human variant of this gene, the *AVPR1A* gene, is important for human behaviors as well. Specifically, there are reports on associations between different variants of the *AVPR1A* gene and such human traits/behaviors as fat intake (Enhorning et al., 2009), personality traits such as altruism (Israel et al., 2008; Prichard, Mackinnon, Jorm, & Easteal, 2007) and aggression (Caldwell, Wersinger, & Young, 2008), and pair bonding (Walum et al., 2008).

To illustrate this last observation, the researchers asked members of 552 Swedish twin pairs, all of whom were living at a given time with a partner, to answer a brief self-report questionnaire with items targeting partner bonding, marital status, and marital problems. All participating twins were genotyped for the 51-microsatellite of the *AVPR1A* gene. It was reported that a particular allele of this polymorphism (the allele RS3 334) was associated with significantly lower scores on the partner-bonding items. This association was true for males only, so that males who were homozygous for this allele were twice as likely to have experienced marital problems or threat of divorce and half as likely to be married if involved in a committed relationship. Moreover, the presence of this allele in the male partner was reported to be correlated with reports of the quality of the relationship's quality as reported by the female partner (Walum et al., 2008).

The second illustration is related to the concept of endophenotypes, which is widely used in the work by Gottesman and colleagues (Gottesman & Gould, 2003) and their many followers. The concept was initially proposed about 35 years ago by Gottesman and Shields (Gottesman & Shields, 1972) and was defined later as a measurable component "unseen by the unaided eye along the pathway between disease and distal genotype" (Gottesman & Gould, 2003, p. 636). When this concept was introduced, the intention, according to the authors, was to fill the gap between the holistic manifestation of the disorders and the gene or genes that govern them. According to the argument for this concept, there is a strong association between the number of genes affecting specific disorders and the number of relevant endophenotypes.[1] In other words, the fewer the genes involved in the manifestation of a particular disorder, the fewer the componential dimensions (i.e., endophenotypes) required to describe and characterize the phenotype; the more genes involved and the more complex the genetic mechanism, the more complex the phenotypic manifestations of the phenotype. From this point of view, the impact of a single gene might be adequately captured by a single dimension of the holistic phenotype, but more dimensions would be required to reflect differentially the impact of multiple genes. Logically, then, it is often assumed that the genetic mechanisms controlling an endophenotype might be simpler for such an endophenotype than for the disorder (or a more complex holistic trait) with which this endophenotype is associated. However, at this point, there is growing evidence that these assumptions might not be quite true.

For example, in the field of intelligence, it has been assumed that the holistic trait of intelligence can be captured by a variety of endophenotypes. Plausible endophenotypes are chronometric indicators such as reaction time, inspection time, and so forth, which are often assumed to be examples of the lower-level cognitive processes that contribute to the complex texture of the *g*-factor of intelligence. It has been argued that the "location" of these phenotypes, being "closer" to the genome than the trait of intelligence itself, should mean that they are governed by less complex genetic mechanisms than those controlling intelligence. Correspondingly, having understood the nature of these genetic mechanisms, the field would have been closer to understanding the genetic nature of intelligence itself. Yet various quantitative–genetic studies indicate that, if anything, there appear to be more genes involved in the control of the variability in chronometric indicators of cognitive processing than for more complex, higher-order cognitive traits. Thus, at least with regard to the number of genes involved in the manifestation of the so-called endophenotypes for intelligence as its chronometric indicators, the assumption that less complex traits are governed by less complex mechanisms might not be either true or helpful in understanding the ways the dynamic genome exerts its influences on cognition.

These observations might also be directly related to classroom practices by informing the differentiation (i.e., lower- vs. higher-level processes) of pedagogical intervention targets.

PSYCHOLOGICAL, EDUCATIONAL, AND GENOMIC SCIENCES: AN EMERGING PLATFORM FOR INTERPRETING INDIVIDUAL DIFFERENCES

Four observations connecting the dynamic genome with complex human behaviors have been made. First, it appears that the different regions of the human genome appear to evolve at different rates; it is assumed that rapidly evolving regions of the dynamic genome might be responsible for the acquisition of specific human traits. Second, although the assumption of the field that the genome controls the brain, which then controls human behavior, remains dominant in the field, the specifics of these connections are rather complex, nonlinear, and perhaps not sequential. These connections reflect an emergence of a true complex dynamic system, where it might be the case that, although the lower levels of the system might arise first developmentally, it is the higher levels of the system that are more constrained genetically. Third, the current evidence indicates that human cognitive functions are not rooted in a particular spot of the brain but rather are routed through its multiple regions. This routing underlies the formation of diverse information processing networks that are amenable to modification through teaching and learning. Fourth, we observed that the complexity of a cognitive process does not necessarily correlate with the complexity of the corresponding genetic mechanism; in fact, more complex cognitive functions can be governed by simpler genetic mechanisms, and simpler genetic mechanisms might be controlled by more complex genetic architecture. Thus, the connections between the dynamic genome and human behaviors are diverse: they range from those that form the foundations for uniquely human behaviors to those that underlie the architecture of cognitive traits.

In concluding this discussion, we point to some of the numerous junctions between the dynamic genome and behavior that are of interest to the fields of psychology and education.

First, development is characterized by a tremendous amount of interindividual variation, which forms the foundation for individual differences. Numerous studies of a variety of human traits, typical and atypical, that have been carried out with genetically informed designs (i.e., designs including pairs or larger groups of genetically related individuals) have indisputably demonstrated that the genome is a major source of these differences. Many researchers are currently working on translating the structural variation in the genome into potentials for individual differences in human behavior in general and cognition in particular. These translations are unfolding in multiple directions, including (a) identifying structural variation in the human genome in specific genes whose proteins can be targeted by pharmacological agents for the purposes of both enhancement and prevention of deterioration of cognitive functioning; (b) understanding the degree of malleability of particular complex behaviors whose manifestation is, at least partially, controlled by genes, especially in response to pedagogical interventions; and (c) developing diagnostic procedures based on identifying genetic risk factors for developmental disorders that present particular challenges for regular classrooms and call for the development and implementation of special pedagogies.

Second, human behavior is governed by the brain in general and by its specific circuits in particular. Many of these specific circuits have now been delineated, at least in broad strokes, allowing researchers to start filling in the specifics of connections between the dynamic genome and these circuits. Yet the field is at its very beginning with regard to its attempts at translating these associations into targeted clinical and educational practices. Much more work is needed to identify the specifics of the dynamics of corticogenesis, the genes that are involved in this process, and the elements in the chain connecting the dynamic genome and the brain that are open to both evolution and targeted, human-orchestrated influences.

Third, the specificity of brain circuits is established anatomically by particular patterns of signal transduction and neuromodulation that develop functionally, in response to environmental stimulation. All these "contributors" to specificity and plasticity are, in turn, governed or assured by specific proteins synthesized by specific genes. Understanding what these genes are and how they function is of direct importance for understanding the regulation of human behavior by the brain. Moreover, such understanding will result in establishing both the sequence and the structure of higher-order functions and inform education with regard to the development of pedagogies best suited to these sequences.

Fourth, brain circuits are responsible for processing and merging information from stimuli that are internal and external to an individual. These stimuli form streams of information that are captured and transduced by various sensory pathways. Identification of the genes engaged in the formation of the biological foundation for these sensory pathways will be important for understanding the connection between the genome and information processing. Having this fundamental bit of understanding in place will inevitably help the field link the dynamic genome to the elusive texture of higher-order human functions such as thought and love. This, of course, will also have direct implications for education, indicating particularly effective strategies for teaching skills that are highly demanded by modern society, such as creativity and compassion.

Clearly this list is not exhaustive and can only grow as the field enhances its understanding of the structure and function of the dynamic genome. Yet in going through this list and ticking off its items, the field formulates a new conceptual framework for further studies of the links between the genome and behavior. This framework brings us far beyond the nature–nurture controversy and opens new and exciting perspectives in understanding human complexity, from its evolutionary roots to its manifestations in the future. And among the many *how* questions there is one that is especially important—how to summarize this wealth of information, benefit from it, and bring it to school with the purpose of further understanding the nature and degree of malleability of individual differences in the classroom.

NOTE

1. A concept utilized to fill the gap between a clinical disorder phenotype and the genome. Other similar concepts are intermediate phenotype, subphenotype, biological marker, subclinical trait, vulnerability marker, and cognitive marker.

AUTHOR NOTE

Preparation of this chapter was supported in part by the following research grants from the National Institutes of Health: R01 DC007665 and P5O HD052120. Grantees undertaking such projects are encouraged to express their professional judgment freely. Therefore this chapter does not necessarily reflect the position or policies of the National Institutes of Health, and no official endorsement should be inferred. The authors are thankful to Ms. Mei Tan for her editorial assistance.

REFERENCES

Baare, W. F., Hulshoff Pol, H. E., Boomsma, D. I., Posthuma, D., de Geus, E. J., Schnack, H. G., et al. (2001). Quantitative genetic modeling of variation in human brain morphology. *Cerebral Cortex, 11,* 816–824.

Barbujani, G., & Pilastro, A. (1993). Genetic evidence on origin and dispersal of human populations speaking languages of the Nostratic macrofamily. *Proceedings of the National Academy of Sciences of the United States of America, 90,* 4670–4673.

Barbujani, G., & Sokal, R. R. (1990). Zones of sharp genetic change in Europe are also linguistic boundaries. *Proceedings of the National Academy of Sciences of the United States of America, 87,* 1816–1819.

Blokland, G. A., McMahon, K. L., Hoffman, J., Zhu, G., Meredith, M., Martin, N. G., et al. (2008). Quantifying the heritability of task-related brain activation and performance during the N-back working memory task: a twin fMRI study. *Biological Psychology, 79,* 70–79.

Butcher, L. M., Davis, O. S. P., Craig, I. W., & Plomin, R. (2008). Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. *Genes, Brain and Behavior, 7,* 435–446.

Buyske, S., Bates, M. E., Gharani, N., Matise, T. C., Tischfield, J. A., & Manowitz, P. (2006). Cognitive traits link to human chromosomal regions. *Behavior Genetics, 36,* 65–76.

Caldwell, H. K., Wersinger, S. R., & Young, W. S. R. (2008). The role of the vasopressin 1b receptor in aggression and other social behaviours. *Progress in Brain Research, 170,* 65–72.

Cavalli-Sforza, L. L., Piazza, A., Menozzi, P., & Mountain, J. (1988). Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Sciences of the United States of America, 85,* 6002–6006.

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

Christiansen, M. H., Kelsey, K., & Tomblin, J. B. (2008). *Associations of ASPM genetic markers with language and non-verbal intelligence*. Paper presented at the BULLD.

Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 10944–10949.

Dick, D. M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., et al. (2006). Linkage analyses of IQ in the collaborative study on the genetics of alcoholism (COGA) sample. *Behavior Genetics, 36,* 77–86.

Donaldson, Z. R., & Young, L. J. (2008). Oxytocin, vasopressin, and the neurogenetics of sociality. *Science, 322,* 900–904.

Dupanloup de Ceuninck, I., Schneider, S., Langaney, A., & Excoffier, L. (2000). Inferring the impact of linguistic boundaries on population differentiation: Application to the Afro-Asiatic-Indo-European case. *European Journal of Human Genetics, 8,* 750–756.

Enhorning, S., Leosdottir, M., Wallstrom, P., Gullberg, B., Berglund, G., Wirfalt, E., et al. (2009). Relation between human vasopressin 1a gene variance, fat intake, and diabetes. *American Journal of Clinical Nutrition, 89,* 400–406.

Evans, P. D., Gilbert, S. L., Mekel-Bobrov, N., Vallender, E. J., Anderson, J. R., Vaez-Azizi, L. M., et al. (2005). Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science, 309,* 1717–1720.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics, 131,* 479–491.

Galton, F. (1869). *Hereditary genius. An inquiry into its laws and consequences.* London, England: Macmillan.

Geschwind, D. H., Miller, B. L., DeCarli, C., & Carmelli, D. (2002). Heritability of lobar brain volumes in twins supports genetic models of cerebral laterality and handedness. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 3176–3181.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 160*, 636–645.

Gottesman, I. I., & Shields, J. (1972). *Schizophrenia and genetics: A twin study vantage point*. New York: Academic Press.

Hulshoff Pol, H. E., Schnack, H. G., Posthuma, D., Mandl, R. C., Baare, W. F., van Oel, C., et al. (2004). Genetic contributions to human brain morphology and intelligence. *Journal of Neuroscience, 26,* 10235–10242.

Huttenlocher, P. R. (1991). Dendritic and synaptic pathology in mental retardation. *Pediatric Neurology, 7,* 79–85.

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Reibold, M., Bachner-Melman, R., et al. (2008). Molecular genetic studies of the arginine vasopressin 1a receptor (AVPR1a) and the oxytocin receptor (OXTR) in human behaviour: From autism to altruism with some notes in between. *Progress in Brain Research, 170,* 435–449.

Jacobs, B., Schall, M., & Scheibel, A. B. (1993). A quantitative dendritic analysis of Wernicke's area in humans. II. Gender, hemispheric, and environmental factors. *Journal of Comparative Neurology, 327,* 97–111.

Jacobs, B., & Scheibel, A. B. (1993). A quantitative dendritic analysis of Wernicke's area in humans. I. Lifespan changes. *Journal of Comparative Neurology, 327,* 83–96.

Kleim, J. A., Lussnig, E., Schwarz, E. R., Comery, T. A., & Greenough, W. T. (1996). Synaptogenesis and Fos expression in the motor cortex of the adult rat after motor skill learning. *Journal of Neuroscience, 16,* 4529–4535.

Koten, J. W. J., Wood, G., Hagroot, P., Goebel, R., Propping, P., Willmes, K., et al. (2009). Genetic contribution to variation in cognitive function: An fMRI study in twins. *Science, 323,* 1737–1740.

Lenroot, R. K., Schmitt, J. E., Ordaz, S. J., Wallace, G. L., Neale, M. C., Lerch, J. P., et al. (2009). Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. *Human Brain Mapping, 30,* 163–174.

Luciano, M., Wright, M. J., Duffy, D. L., Wainwright, M. A., Zhu, G., Evans, D. M., et al. (2006). Genome-wide scan of IQ finds significant linkage to a quantitative trait locus on 2q. *Behavior Genetics, 36,* 45–55.

Mandelman, S. M., & Grigorenko, E. L. (in press). Intelligence: Genes, environments, and everything in between. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence*. New York: Cambridge University Press.

Matthews, S. C., Simmons, A. N., Strigo, I., Jang, K., Stein, M. B., & Paulus, M. P. (2007). Heritability of anterior cingulate response to conflict: an fMRI study in female twins. *Neuroimage, 38,* 223–227.

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P., et al. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics, 9,* 356–369.

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.

McMahon, R. (2004). Genes and languages. *Community Genetics, 7,* 1–13.

Mekel-Bobrov, N., Gilbert, S. L., Evans, P. D., Vallender, E. J., Anderson, J. R., Hudson, R. R., et al. (2005). Ongoing adaptive evolution of ASPM, a brain size determinant in Homo sapiens. *Science, 309,* 1720–1722.

Mekel-Bobrov, N., Posthuma, D., Gilbert, S. L., Lind, P., Gosso, M. F., Luciano, M., et al. (2007). The ongoing adaptive evolution of ASPM and Microcephalin is not explained by increased intelligence. *Human Molecular Genetics, 16,* 600–608.

Miller, G. F., & Penke, L. (2007). The evolution of human intelligence and the coefficient of additive genetic variance in human brain size. *Intelligence, 35,* 97–114.

Nettle, D., & Harriss, L. (2003). Genetic and linguistic affinities between human populations in Eurasia and West Africa. *Human Biology, 75*, 331–344.

Polk, T. A., Park, J., Smith, M. R., & Park, D. C. (2007). Nature versus nurture in ventral visual cortex: A functional magnetic resonance imaging study of twins. *Journal of Neuroscience, 27,* 13921–13925.

Poloni, E. S., Semino, O., Passarino, G., Santachiara-Benerecetti, A. S., Dupanloup, I., Langaney, A., et al. (1997). Human genetic affinities for Y-chromosome P49a,f/TaqI haplotypes show strong correspondence with linguistics. *American Journal of Human Genetics, 61,* 1015–1035.

Posthuma, D., De Geus, E. J., Baare, W. F., Hulshoff Pol, H. E., Kahn, R. S., & Boomsma, D. I. (2002). The association between brain volume and intelligence is of genetic origin. *Nature Neuroscience, 5,* 83–84.

Posthuma, D., de Geus, E. J. C., & Boomsma, D. I. (2003). Genetic contributions to anatomical, behavioral, and neurophysiological indices of cognition. In R. Plomin, J. C. DeFries, I. W. Craig, & P. McGuffin (Eds.), *Behavioral genetics in the postgenomic era* (pp. 141–161). Washington, DC: American Psychological Association.

Posthuma, D., Luciano, M., Geus, E. J., Wright, M. J., Slagboom, P. E., Montgomery, G. W., et al. (2005). A genomewide scan for intelligence identifies quantitative trait loci on 2q and 6p. *American Journal of Human Genetics, 77,* 318–326.

Prabhakar, S., Visel, A., Akiyama, J. A., Shoukry, M., Lewis, K. D., Holt, A., et al. (2008). Human-specific gain of function in a developmental enhancer. *Science, 321,* 1346–1350.

Prichard, Z. M., Mackinnon, A. J., Jorm, A. F., & Easteal, S. (2007). AVPR1A and OXTR polymorphisms are associated with sexual and reproductive behavioral phenotypes in humans. Mutation in brief no. 981. Online. *Human Mutation, 28,* 1150.

Robinson, G. E. (2004). Genomics. Beyond nature and nurture. *Science, 304,* 397–399.

Robinson, G. E., Fernald, R. D., & Clayton, D. F. (2008). Genes and social behavior. *Science, 322,* 896–900.

Rosser, Z. H., Zerjal, T., Hurles, M. E., Adojaan, M., Alavantic, D., Amorim, A., et al. (2000). Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *American Journal of Human Genetics, 67,* 1526–1543.

Sakai, K. L., Miura, K., Narafu, N., & Muraishi, Y. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex, 14,* 1233–1239.

Shaw, P. (2007). Intelligence and the developing human brain. *Bioessays, 29,* 962–973.

Sommer, I. E., Ramsey, N. F., Mandl, R. C., & Kahn, R. S. (2002). Language lateralization in monozygotic twin pairs concordant and discordant for handedness. *Brain, 125,* 2710–2718.

Toga, A. W., & Thompson, P. M. (2005). Genetics of brain structure and intelligence. *Annual Review of Neuroscience, 28,* 1–23.

Wainwright, M. A., Wright, M. J., Luciano, M., Montgomery, G. W., Geffen, G. M., & Martin, N. G. (2006). A linkage study of academic skills defined by the Queensland Core Skills Test. *Behavior Genetics, 36,* 56–64.

Wallace, G. L., Eric Schmitt, J., Lenroot, R., Viding, E., Ordaz, S., Rosenthal, M. A., et al. (2006). A pediatric twin study of brain morphometry. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 47,* 987–993.

Walum, H., Westberg, L., Henningsson, S., Neiderhiser, J. M., Reiss, D., Igl, W., et al. (2008). Genetic variation in the vasopressin receptor 1a gene (AVPR1A) associates with pair-bonding behavior in humans. *Proceedings of the National Academy of Sciences of the United States of America, 1105,* 14153–14156.
Winterer, G., & Goldman, D. (2003). Genetics of human prefrontal function. *Brain Research Reviews, 43,* 134–163.
Wright, I. C., Sham, P., Murray, R. M., Weinberger, D. R., & Bullmore, E. T. (2002). Genetic contributions to regional variability in human brain structure: Methods and preliminary results. *Neuroimage, 17,* 256–271.

2 Dyslexia Deciphered

RICHARD K. WAGNER AND PATRICIA T. KANTOR

Most children can to learn to read with effective instruction. But between 3% and 20% of children, depending on the specific criteria used, fail to learn to read adequately. Roughly 6%, or nearly 3 million children, received assistance for specific learning disabilities from federally funded programs in public schools in the 2003–2004 school year, the most recent year for which statistics are available (National Center for Educational Statistics, 2006). This percentage is up from less than 2% in the 1976–1977 school year. The vast majority of students who receive assistance for specific learning disabilities have problems in reading.

The consequences of failing to learn to read adequately are sobering. Long-term longitudinal studies that follow students over many years indicate that children who are behind in reading at the end of first grade rarely catch up with the peers (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Torgesen & Burgess, 1998; Torgesen, Rashotte, & Alexander, 2001). Poor reading in third grade is a disconcertingly accurate predictor of dropping out of school (National Institute of Child Health and Human Development, 2000).

Support for this research has been provided by Grant P50 HD052120 from NICHD.

What do we know about the causes of dyslexia, the most pronounced and persistent example of failing to learn to read adequately? What have we discovered about how individuals with dyslexia can be identified so that they can be given help? The purpose of this chapter is to consider some important discoveries about the nature of failing to learn to read and how they are being applied.

EARLY VIEWS OF THE NATURE OF DYSLEXIA

The first description of dyslexia was provided by the Scottish surgeon James Hinshelwood in 1905. He labeled it congenital word blindness, noting that it tended to run in families and was characterized by an inability to read words on a page despite apparently adequate vision otherwise.

A widespread early belief about dyslexia was that it resulted from seeing words backward, and this remains the most popular view. Individuals with dyslexia are reported to read WAS as SAW, or to confuse the letters b and d. A possible explanation of this phenomenon was subsequently provided by Samuel Orton's theory of mixed cerebral dominance (Hallahan & Mock, 2003). As infants develop, responsibility for performance in specific domains tends to become more localized in one hemisphere of the brain. Being right- or left-handed is one manifestation of cerebral dominance. According to Orton's theory, what the left hemisphere views as a d is viewed by the right hemisphere as a b. For normal readers, the normal dominance of one of the hemispheres precludes a conflict in perception. However, for individuals with dyslexia, neither hemisphere is dominant, resulting in conflicting perceptions.

Orton's theory was eventually supplanted by a relatively simple explanation of reversal errors. Children with reading problems rarely stand out before the end of first grade or in second grade. They do indeed make reversal errors, confusing *was* with *saw* and *b* with *d*. But it turns out that these kinds of reversal errors are among the easiest to make; they are quite common in kindergarten and in early first-grade classrooms among normally developing readers. Reading *was* as *was* rather than *saw* requires learning to read from left to right, which is an arbitrary direction for traversing a page of print. Other writing systems are characterized by reading from right to left or even from top to bottom.

To confuse letters such as *b* and *d* is understandable, as they are similar both visually and in sound. Both are called stop consonants because the sounds of *b* and *d* are made by obstructing or stopping the flow of air in the vocal tract and then releasing it abruptly.

It turns out that poor readers in second grade do not make proportionately more reversal errors than do younger normal readers who are matched at the same level of reading (Crowder & Wagner, 1992; Werker, Bryson, & Wassenberg, 1989). The reversal errors of older poor readers merely reflect the fact that they are behind in reading; it does not explain their poor reading. The association of reversals with poor reading probably results from the fact that in second grade and beyond, only poor readers are still making reversal errors; consequently they stand out from their classmates.

Another popular view about the nature of dyslexia was that it is caused by deficient eye movements. Reading requires remarkably sophisticated and accurate eye movements (Rayner & Pollatsek, 1989). As you read these words, you probably feel that your eyes are gliding smoothly across the page. This perception turns out to be wrong. A simple experiment will demonstrate this. Ask a friend to read directly across from you, holding the reading material low enough so that you can observe your friend's eyes while he or she reads. If you look carefully, you will see that your friend's eyes move in a series of small but observable jerky movements. The eyes actually are propelled across the page in a series of tiny ballistic movements called saccades. The movements are ballistic much like those of a cannon ball that is driven by an initial burst of energy. During the saccades, little information is available to the eyes because they are moving too fast to see letters or words clearly. During fixations or pauses between the saccades, the eyes are able to perceive letters and words.

The belief that faulty eye movements are the cause of the dyslexia is understandable if you repeat the same experiment with an individual who reads poorly. What you will observe is that this person's eye movements are noticeably more erratic.

The view that dyslexia is caused by faulty eye movements has been challenged by studies that manipulated difficulty of the reading material for good and poor readers. In these studies, normal readers were given material that was as difficult for them to read as is grade-level reading material for individuals with reading difficulties. Under these conditions, the eye movements of normal readers resembled the more erratic eye movements of poor readers. Conversely, when poor readers were given easy reading material that they were able to read as well as normal readers read grade-level text, the eye movements of poor readers now resembled those of good readers (Crowder & Wagner, 1992). These results suggest that faulty eye movements are not the cause of poor reading but a by-product of it. If you cannot read the words on the page, your eyes will not progress across the page in the same way they would if you could read the words.

CURRENT VIEWS REGARDING THE NATURE OF DYSLEXIA

Hinshelwood was correct about two aspects of dyslexia. It does run in families and it is characterized by poor performance at reading the words on a page (Adams, 1990; National Research Council, 1998; Stanovich, 1982; Vellutino, 1979). Although many individuals with dyslexia are also impaired in reading comprehension, the impaired comprehension is often a downstream consequence of a primary impairment in word-level reading (Aaron, 1989; Stanovich & Siegel, 1994).

The word-level reading problem, in turn, appears to derive from a problem in language rather than vision. To understand how dyslexia can result from a problem in language rather than in vision requires knowledge of how language and print are related.

Linking Language and Print

Writing systems have been designed to represent various aspects of their corresponding spoken language (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). A beginning reader's task is to figure out the nature of this correspondence. Doing so successfully involves relating aspects of phonology, morphology, and orthography.

Phonology and Morphology

Speech is represented at multiple, hierarchical linguistic levels that connect phonological and morphological aspects of language. Beginning at the lowest level of the hierarchy and proceeding to the top, four linguistic levels can be distinguished. These are the acoustic, phonetic, phonological, and morphophonological levels (Crowder & Wagner, 1992).

Speech is represented at the *acoustic* level by nearly continuous waves of acoustic energy, which can be observed on a spectrogram. A spectrogram displays acoustic energy by frequency over time. If you observe a spectrogram of a spoken sentence such as "The rain in Spain falls mainly in the plain," you will, quite remarkably, see virtually no spacing or breaks. Even though we easily distinguish individual words, syllables within words, and sounds within syllables, the spectrogram reveals nearly continuous energy. Our ability to perceive distinct words and parts of words is due to our perceptual and cognitive machinery as opposed to a characteristic inherent in the physical signal itself. This explains why speakers of languages you do not know appear to speak very rapidly and the words seem to run together.

Speech is represented at the *phonetic* level by phones. *Phones* belong to the universe of individual sounds made by all speakers. At the phonetic level, the sounds of the p in top, spot, and pot are represented by three distinct phones. You can confirm this assertion with a simple experiment. Hold your hand in front of your mouth while you say the three words *top, pot,* and *spot*. You will feel a relatively strong burst of air while pronouncing the sound of the *p* in *top;* a somewhat less strong burst of air while pronouncing the sound of the *p* in *pot;* and very little air movement while pronouncing the sound of the *p* in *spot*. The differences you feel confirm that the three sounds are pronounced differently.

Speech is represented at the *phonological* level by abstract units call phonemes. *Phonemes* are sound distinctions that signal changes in meaning in a given language. The sounds represented by the *f* in *fit* and the *p* in *pit* are different phonemes, which signals the fact that these two speech segments are associated with different meanings. Phonemes consist of families of phones. Thus, the sounds of the *p* in *top*, *spot*, and *pot* all are representations of the /p/ phoneme and are referred to as allophones of the phoneme /p/. What is abstract about the nature of phonemes is that they are categories of phones that signal meaning in a given oral language. In contrast, the concrete distinctions represented by phones can be observed directly, as in the example of pronouncing words while holding your hand in front of your mouth.

Speech is represented at the top *morphophonemic* level by strings of phonemes that represent morphemes or units of meaning. *Morphology* refers to the composition of a word with respect to morphemes or "minimal meaningful elements" (Bloomfield, 1933). Strings of phonemes that represent units of meaning are called *morphophonemes*. Analogous to phonemes consisting of a family of allophones, morphemes consist of families of allomorphs. For example, the letter string *sign* is a morpheme that is represented in the allomorphs in the printed words *signed* and *signature*. The English writing system is morphophonemic in that pronunciations are represented for the most part, but with compromises so as to convey meaning. This is reflected in the fact that written words *signed* and *signature* share the spelling *sign* even though the pronunciations of the *sign* parts of both words are remarkably different.

Orthography

Orthography refers to the system of marks used to represent pronunciation and meaning in writing. For example, English orthography consists of the 26 upper- and lowercase letters, the numerals 0 through 9, punctuation marks, and a small number of other symbols (e.g., $, #, and @). Looking around the world today, there are many writing systems. However, they can be categorized into three basic types (Crowder & Wagner, 1992; DeFrancis, 1989, Gelb, 1952; Rayner et al., 2001; Rayner & Pollatsek, 1989). They are *alphabets, syllabaries,* and *morphosyllabaries.*

Alphabetic writing systems—such as Dutch, English, German, Italian, Korean, Spanish, and the Chinese pin yan system—rely on a relatively small number of letters that correspond roughly to the phonological level of speech representation. Letters represent phonemes rather than phones or morphophonemes. In the examples of *top, spot,* and *pot,* the three phones represented by the sounds associated with the letter *p* were allophones of the single phoneme /p/, and the single letter *p* was used in their spellings. Syllabaries are writing systems with characters that represent syllables rather than phonemes. An example of a syllabary is the Japanese Kana. Finally, morphosyllabic writing systems are writing systems with characters that represent not only syllables but also morphemes or meaning units. Examples of morphosyllabic writing systems include Chinese and the Japanese Kanji system.

The Development of Phonological Awareness in Learning to Read English

Phonological awareness refers to an awareness of and access to the sound structure of one's oral language (Jorm & Share, 1983; Wagner & Torgesen, 1987). All spoken English words of midwestern American English can be represented using the phonemes presented in Table 2.1. (*Midwestern American English* refers to the dialect typically spoken by news anchors on American television.) Vowels are produced by air passing over the vocal cords, which causes them to vibrate. Consonants are produced by obstructing and releasing air. Which consonant is produced depends upon which parts of the vocal tract are used to control the flow of air (e.g., teeth, lips, palate) and whether or not the vocal cords are also vibrating.

Table 2.1

PHONEMES OF MIDWESTERN AMERICAN ENGLISH

CONSONANTS								
Stops	Voiced	**b**arn	**d**own	**g**ive				
	Voiceless	**p**ickle	**t**ell	**c**art				
Fricatives	Voiced	ha**v**e	**th**eer	**z**oo	trea**s**ure			
	Voiceless	**f**ickle	**th**ink	**s**oup	**sh**ip	**h**at		
Affricatives	Voiced	**j**ustice						
	Voiceless	**ch**art						
Nasals		**m**ill	**n**ose	thi**ng**				
Liquids		**l**ight	**r**ed					
Glides		**y**acht	**w**orld					
VOWELS								
Front		z**e**bra	m**i**t	b**e**t	s**a**t			
Middle		b**er**m	c**u**p	**a**bout	h**oo**d			
Back		st**ar**t	n**o**d	st**or**k	d**oo**dle			
Dipthongs		c**ow**	m**y**	fl**oa**t	k**oi**	d**ee**r	p**oo**r	f**er**ret
			sn**a**ke					

There are approximately 10 trillion possible combinations of the phonemes presented in Table 2.1. Of these, only a relatively small number actually occur in spoken language, and many of these occur in multiple words (Wagner et al., 1997). Thus, *sit,' knit,* and *fit* each consist of three phonemes, the first of which is different and the latter two are identical in the three words. These facts are represented by their spellings. They have different initial letters and identical medial and final letters. The correspondence between phonemes and letters suggests that an awareness of the phonological nature of one's oral language might make it easier to learn to read.

Two kinds of evidence suggest that phonological awareness indeed plays a causal role in the acquisition of beginning reading. The first kind of evidence is provided by longitudinal correlational studies. The second kind is provided by training or interventional studies.

Consider an example of causal modeling of data collected in a longitudinal study of the development of phonological awareness and learning to read. Wagner et al. (1997) reported a 5-year longitudinal study of 216 children who were followed from kindergarten through fourth grade. To minimize measurement error and get the best possible view of developing relations, they analyzed latent variables. Latent variables represent constructs like phonological awareness with multiple measures. A latent variable captures the common variance among its multiple measures. Measurement error is minimized because it is specific to measures rather than common across measures.

Three measures of phonological awareness were used in this study. The first measure, called elision, requires dropping out part of a spoken word ("Say the word *task.* Now tell me what word would be left if I said *task* without saying /t/"). The second measure was sound categorization, a task requiring picking the odd one out in terms of sound of a set of items (*bit, knit, mat, sit*). The third task was phoneme segmentation, a task that involved listening to a word and then saying the individual sounds it contains.

In addition to phonological awareness, the study included measures of phonological memory and rapid naming. *Phonological memory* refers to coding information into a sound-based representation system for temporary storage (Baddeley, 1982, 1986; Conrad, 1964). The two measures of phonological memory included in the study were digit span and memory for sentences. Finally, rapid naming was included as a measure of the efficiency with which phonological codes can be retrieved from permanent memory. The measures of rapid naming used in the study required that children name series of digits or letters as rapidly as they could.

Control variables included vocabulary and the autoregressive effect of word-level reading at the previous time point. The autoregressor was included to rule out finding a spurious causal relation. For example, it might be the case that the only causal variable operating on the outcome of word-level reading measured in third grade is word-level reading in first grade. In other words, if you already do well in reading in first grade, you are likely to continue to do well because you are starting with an advantage. If the autoregressor of first-grade reading is omitted from the model, any variable included in the model that is correlated with the omitted first-grade reading variable will show a spurious causal influence by virtue of its correlation with the omitted true cause.

The key results for determining whether phonological awareness plays a causal role in learning to read are presented in Table 2.2. The

Table 2.2

INDEPENDENT CAUSAL INFLUENCES OF PHONOLOGICAL PROCESSING AND CONTROL VARIABLES ON SUBSEQUENT WORD-LEVEL READING

	TIME PERIOD		
EXOGENOUS VARIABLE	K TO 2ND GRADE	1ST TO 3RD GRADE	2ND TO 4TH GRADE
Phonological processing latent variables			
Phonological awareness	.37***	.29*	.27***
Phonological memory	.12	–.03	.07
Rapid naming	.25*	.21*	.07
Control variables			
Vocabulary	.10	.22***	–.01
Autoregressor	.02	.27*	.57***

*$p < .05$. ***$p < .001$.
Adapted from "Changing Relations Between Phonological Processing Abilities and Word-Level Reading as Children Develop From Beginning to Skilled Readers: A 5-Year Longitudinal Study," by R. K. Wagner, J. K. Torgesen, C. A. Rashotte, S. A. Hecht, T. A. Barker, S. R. Burgess, et al., 1997. *Developmental Psychology, 33,* 468–479.

values in the table represent the unique or independent causal influence of the measures of phonological processing and the control variables on subsequent reading. The results were that phonological awareness exerted an independent causal influence on word-level reading at each of the three time intervals examined. The growing influence of the autoregressor reflects the increasing stability of individual differences in word-level reading over time (i.e., the idea that a good reader in first grade is likely to be a good reader in third grade).

It also is possible that a reciprocal causal relation exists if learning to read improves phonological awareness. Results addressing this possibility are presented in Table 2.3. Here the outcome was subsequent phonological awareness and the causal variable was letter knowledge. The results suggest an independent causal influence of letter knowledge on subsequent phonological awareness.

Table 2.3

RECIPROCAL CAUSAL INFLUENCES OF LETTER KNOWLEDGE ON SUBSEQUENT PHONOLOGICAL AWARENESS

	TIME PERIOD	
EXOGENOUS VARIABLE	K TO 2ND GRADE	1ST TO 3RD GRADE
Letter knowledge	.23**	.12*
Control variables		
Vocabulary	.19*	.17**
Autoregressor	.43***	.70***

*$p < .05$. **$p < .01$. ***$p < .001$.
Adapted from "Changing Relations Between Phonological Processing Abilities and Word-Level Reading as Children Develop From Beginning to Skilled Readers: A 5-Year Longitudinal Study," by R. K. Wagner, J. K. Torgesen, C. A. Rashotte, S. A. Hecht, T. A. Barker, S. R. Burgess, et al., 1997. *Developmental Psychology, 33,* 468–479.

These results provide evidence of bidirectional causal influences between developing phonological processing variables and reading: Individual differences in early phonological awareness play a causal role in subsequent individual differences in word-level reading skills, and individual differences in early letter knowledge play a causal role in subsequent individual differences in phonological awareness.

The second kind of evidence relevant to a possible causal effect of phonological skills in learning to read comes from interventional studies. Those studies that have targeted phonological awareness and relations between letters and sounds have produced gains in the decoding skills of poor readers (Adams, 1990; Ball & Blachman, 1991; Brady, Fowler, Stone, & Winbury, 1994; Bus & van Ijzendoorn, 1999; Byrne & Fielding-Barnsley, 1989, 1991, 1993, 1995; Byrne, Fielding-Barnsley, & Ashley, 2000; Chall, 1967/1983; Ehri, Nunes, Stahl, & Willows, 2001; Foorman et al., 2003; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Foorman, Francis, Novy, & Liberman, 1991; Hatcher, Hulme, & Ellis, 1994; Lovett, Steinbach, & Frijters, 2000; National Research Council, 1998; Rashotte, MacPhee, & Torgesen, 2001; Schneider, Ennemoser, Roth, & Kuespert, 1999; Schneider, Roth, & Ennemoser, 2000; Torgesen et al., 2001; Torgesen et al., 1999; Vellutino, Scanlon, & Tanzman, 1998; Wise, Ring, & Olson, 1999).

Some evidence of a causal relation between phonological awareness and learning to read can be challenged on the basis of possible alternative explanations (Castles & Coltheart, 2004). For example, if both phonological awareness and letter knowledge are trained in an intervention, as is commonly done, how can we be sure it is the phonological awareness training that is responsible for a boost in word reading? It is possible that the combination of training phonological awareness and letter knowledge actually is a proxy for phonics instruction, something we already know to be effective. Of course, a proponent of a causal relation between phonological awareness and learning to read might argue that one reason for the effectiveness of phonics instruction is that it produces a by-product of improved phonological awareness. Although individual studies can be challenged on various grounds, the overall pattern of results, and in particular the longitudinal correlational studies that have included an autoregressor effect, support the existence of a causal influence of phonological awareness on word reading (Hulme, Snowling, Caravolas, & Carroll, 2005).

Implications of the Phonological Nature of Dyslexia

One practical application that has emerged from research on relations between phonological processing and reading has been the development of tests of phonological processing for use in identifying individuals with dyslexia. An example of such a test is the Comprehensive Test of Phonological Processing (CTOPP); (Wagner, Torgesen, & Rashotte, 1999).

The CTOPP was constructed to measure phonological awareness, phonological memory, and rapid naming for individuals ranging in age from 5 to 24. Accommodating this wide age range and the remarkable development of phonological processing that occurs during it required developing two versions of the test. The first version is for individuals who are 5 or 6 years old and consists of seven core subtests and one supplemental subtest. The second version, for ages 7 through 24, consists of six core subtests and six supplemental subtests. The supplemental subtests are provided to permit more in-depth assessment of strengths and weaknesses in phonological processing. Both versions are individually administered, and the core subtests take roughly 30 minutes to administer.

The 13 subtests included on one or both versions of the CTOPP as either a core or supplemental subtest are listed in Table 2.4. The phonological awareness subtests assess skills that are relevant to reading instruction. Elision requires segmenting spoken words into smaller parts that correspond to spellings. Segmenting Words assesses a skill that is prerequisite to letter-sound mapping, in that letter-sound mapping requires that words be segmented into letter-size parts. Blending Words requires combining sounds together, and is required when decoding words letter by letter or by larger phonological units and then combining the decoded parts to identify the target word. Sound Matching requires identification of the first sound in a word; this is one of the first reading skills taught.

Table 2.4

THE SUBTESTS OF THE COMPREHENSIVE TEST OF PHONOLOGICAL PROCESSING (CTOPP)

CORE PHONOLOGICAL AWARENESS SUBTESTS

1. *Elision*. This is a 20-item subtest that measures the ability to drop a phonological segment from a word and pronounce the remaining portion. For example, an examinee might be asked to say the word *cat* without saying /c/. The correct answer is *at.*

2. *Blending Words*. This is a 20-item subtest that measures the ability to combine phonological segments into words. For examine, an examinee might be asked, "What word do these sounds make: /c/ /a/ /k/?" The correct answer is *cake*.

3. *Sound Matching*. This 20-item subtest measures the ability to match sounds. The examiner pronounces a word and then three responses. All words are represented by pictures. For the first half of the subtest, the task is to match initial sounds. For example, the examinee might be asked, "Which word begins with the same sound as *pan: hat, pig,* or *cone*? The correct answer is *pig.* For the second half of the subtest, the task is to match final sounds, so for the same items, *cone* would now be the correct answer because it ends in the same sound as *pan.*

CORE PHONOLOGICAL MEMORY SUBTESTS

4. *Memory for Digits*. This 21-item subtest measures forward digit span for strings varying in length from 2 to 8 digits in length.

5. *Nonword Repetition*. This is an 18-item subtest that assesses phonological memory by asking individuals to repeat nonwords ranging in length from 3 to 15 phonemes. For example, the examinee might listen to and then pronounce the nonword *jatsiz.*

Table 2.4

THE SUBTESTS OF THE COMPREHENSIVE TEST OF PHONOLOGICAL PROCESSING (CTOPP) ***(Continued)***

CORE RAPID NAMING SUBTESTS

6. *Rapid Color Naming*. This 72-item subtest measures the speed at which a series of pictured colored blocks can be named. It is administered in two trials, each of which requires the individual to name 36 items as quickly and accurately as possible. The colors to be named are blue, red, green, black, yellow, and brown. This is a core subtest for the 5- and 6-year-old version of the CTOPP, and a supplemental subset for the 7- through 24-year-old version.

7. *Rapid Object Naming*. This subtest is identical to Rapid Color Naming except that the stimuli to be named are the objects pencil, star, fish, chair, boat, and key. This also is a core subtest for the 5- and 6-year-old version of the CTOPP, and a supplemental subset for the 7- through 24-year-old version.

8. *Rapid Digit Naming*. This subtest is identical to Rapid Color Naming except that the stimuli to be named are the digits 2,3, 4, 5, 7, and 8. This is a core subtest for the 7- through 24-year-old version only.

9. *Rapid Letter Naming*. This subtest is identical to Rapid Color Naming except that the stimuli to be named are the letters a, c, k, n, s, and t. This is a core subtest for the 7- through 24-year-old version only.

SUPPLEMENTAL SUBTESTS

10. *Phoneme Reversal*. This 18-item subtest requires reversing the order of sounds in a speech segment and pronouncing the resultant word. For example, the examinee might be asked to say “ood” backward. The correct answer is “do.”

11. *Blending Nonwords*. This 18-item subtest is identical to the Blending Words subtest except that the stimuli are nonwords rather than words.

12. *Segmenting Words*. This 20-item subset requires an individual to separate a word into its constituent parts and pronounce the parts individually. For example, the examinee might be asked to say the word “cat” one sound at a time.

13. *Segmenting Nonwords*. This 20-item subtest is identical to the Segmenting Words subtest except that the stimuli are nonwords.

Blending and Segmenting Nonwords are included for clinicians and researchers who prefer nonword items. Some authorities believe that nonword items provide purer measures of phonological processing because the influence of vocabulary is reduced. However, nonword items are more difficult for examiners to administer and score. In addition to presenting problems for some examiners, some examinees, especially younger children, have difficulty working with or pronouncing nonwords. Based on these concerns, nonword items were included only on supplemental subtests. Also, to reduce error attributable to individual differences in the accuracy with which examiners are able to pronounce nonwords, items for subtests involving nonwords are presented via digital recording.

Turning to phonological memory, a number of digit span tasks are available on various tests. Digit Span on the CTOPP was constructed to be a better measure of phonological memory than is provided by the typical digit span task. To maximize the extent to which this subtest assesses phonological memory rather than rehearsal or memory strategies, the digits are presented at a rate of 2 per second. This rate of presentation minimizes the opportunity for rehearsal or use of other memory strategies to affect performance. All items require forward recall of digits only because backward recall introduces cognitive complexity and strategic processing; it may be a better measure of working memory than of phonological memory per se. Two strategies were used to improve reliability. First, the digit span items are presented using a digitally recorded format. This ensures that the digits are presented at a rate of 2 per second and that rate of presentation does not vary across examiners. Second, three items are presented at each string length, and the score is the number of items recalled correctly as opposed to a span length. The items for the Nonword Repetition task were created in a way that minimizes their similarity to actual words. The items were constructed by randomly combining phonemes and then retaining for possible inclusion resultant strings that were pronounceable. Other nonword repetition tasks have items that are more word-like, which encourages lexical processing using knowledge of real words as opposed to phonological processing.

Turning to rapid naming, 5 and 6-year-olds are asked only to name common objects and colors. Although rapid naming of digits and letters is more predictive of reading, younger children are asked only to name objects and colors because rapid naming tasks ought to measure differences in rates for naming familiar items as opposed to differences in familiarity with the items to be named. When young children are asked to name digits and letters, their performance can be affected by their limited knowledge of the digits and letters to be named. Rapid naming of letters and digits forms the core subtests for older children. However, rapid naming of colors and objects is available in supplemental subtests for those interested in a measure of rapid naming that is less affected by letter or digit knowledge.

EARLY IDENTIFICATION OF DYSLEXIA

The CTOPP has been widely used to identify students with dyslexia. However, there is a clear need to identify children at risk for the development of dyslexia well before learning to read has begun (Fletcher, Lyon, Fuchs, & Barnes, 2007). The immediate problem is that preschool-age children are not able to perform phoneme-based phonological awareness tasks, nor are they able to read. A solution has resulted in the identification of emergent forms of both phonological processing and literacy that are predictive of their full-blown counterparts years later (Snow, Burns, & Griffin, 1998; Whitehurst & Lonigan, 1998).

The ability to develop phonological awareness tasks that work for preschool-age children rests on the construct of linguistic complexity. The basic idea is that larger linguistic units are easier to manipulate than are smaller linguistic units (Anthony, Lonigan, Burgess, Driscoll, & Phillips, 2002; Lonigan, 2006). A hierarchy of linguistic units ordered from large to small includes compound words, syllables, onset-rime units within syllables, and finally individual phonemes within onset-rime units. The onset of a syllable refers to the initial consonant or consonant cluster. The rime refers to the vowel and final consonant or consonant cluster. For example, the single-syllable word *cat* consists of the onset in the form of the sound of the *c* and the rime in the form of the sounds of the *at*.

Returning to the task of elision, how difficult an item is depends on the level of linguistic complexity of the unit that must be manipulated (Anthony & Francis, 2005; Anthony, Lonigan, Driscoll, Phillips, &Burgess, 2003). Preschool nonreaders are able to perform elision on compound words (e.g., "Say *starfish*." "Now say *starfish* without saying *fish*."), syllables (e.g., "Say *running*." "Now say *running* without saying *ing*."), and to some extent, on onset-rime units within a syllable (e.g., "Say *rant*." "Now say *rant* without saying /r/."). Although an onset can be an individual phoneme, and it thus can appear that preschool nonreaders are able to access and manipulate individual phonemes, elision tasks involving onset-rime can be performed by manipulating the more accessible rime unit. Preschool nonreaders typically are not able to perform elision on individual phonemes (e.g., "Say *toad*." "Now say *toad* without saying the /d/."), and isolating a phoneme from a cluster of phonemes is more difficult still (e.g., "Say *past*." "Now say *past* without saying the /s/ sound."). A test of phonological awareness that relies on easier levels of linguistic complexity can be found on the Test of Preschool Early Literacy (TOPEL); (Lonigan, Wagner, Torgesen & Rashotte, 2007). It has proven useful with children as young as age 3.

Print awareness refers to an emergent form of reading that can be assessed in preschool-age children that is predictive of later reading (National Early Literacy Panel, 2008; Whitehurst & Lonigan, 1998). The TOPEL subtest Print Knowledge is designed to measure print awareness. It includes knowledge about written conventions and letter knowledge. For example, children are shown three pictures and a word and are asked to indicate which one has letters. They are shown pictures of four books, one of which has a title on the cover, and are asked which one shows the name of a book. Children are shown pictures of writing and asked to point to the one that has the most words. Children are asked to differentiate letters from numbers and other symbols. Finally, children are asked about the names and sounds of letters.

Using tasks such as elision with compound words and print awareness, children at risk for reading failure can be identified as early as 3 years of age. The hope is that interventions designed to improve the development of their phonological abilities can be implemented in preschool to reduce the risk for having subsequent reading problems or, if that cannot be avoided entirely, reducing their severity.

PROBLEMS IN LEARNING TO READ OTHER SCRIPTS

Although much of the research on dyslexia is based on participants learning to read English, a substantial literature now exists on problems learning to read scripts other than English. Three principles emerge from cross-linguistic studies of dyslexia.

First, deficits in phonological processing in general and in phonological awareness in particular appear in every language that has been studied to date (Goswami, 2002; Goulandris, 2003; Ziegler & Goswami, 2005). The relation between phonological awareness and learning to read is not specific to written English: A universal characteristic of poor beginning readers regardless of orthography is a deficit in phonological awareness (Shaywitz, Morris, & Shaywitz, 2008). Although sporadic reports appear to contradict this remarkably general assertion for a given orthography (e.g., Wydell, 2003), these reports routinely are followed up by other studies that provide evidence of a deficit in phonological awareness for poor readers of the target orthography (e.g., Seki, Kassai, Uchiyama, & Koeda, 2008). At least some of the apparently discrepant findings are attributable to differences in cultural practices with respect to the acknowledgment of the existence of individuals who are poor readers and how they are identified and categorized (Tzeng, 2007).

Second, the nature and severity of the impact of a deficit in phonological awareness depends upon the linguistic complexity represented by a given script. All written languages represent spoken languages (Goswami, 2002). The level at which this representation takes place affects the manifestation of a deficit in phonological awareness. Recall the levels of linguistic complexity described previously. Discriminating and manipulating relatively large linguistic units such as syllables is easier than discriminating and manipulating small linguistic units such as phonemes. Consequently, scripts with orthographic elements that correspond to smaller linguistic units, such as phonemes, are more challenging to master for individuals with a deficit in phonological awareness than are scripts that represent larger linguistic units, such as syllables.

Third, within alphabetic scripts, the regularity of the mapping between letters and phonemes affects the developmental course of the impact of a deficit in phonological awareness. For some written languages such as Finnish, a given letter is associated with a single phoneme and all words are "regular" in the sense of being decodable simply by applying the associated phonemes to the letters present in a target word. For this reason, these written languages are characterized as being transparent. For other written languages such as English or French, a given letter can be associated with multiple phonemes and a substantial proportion of words are irregular (e.g., the written word *yacht*). These written languages are characterized as being opaque. A deficit in phonological awareness affects learning opaque languages over a longer period of time than it does learning transparent languages. For example, Georgiou, Parrila, and Papadopoulos (2008) compared young children learning to read the transparent written language of Greek with children learning to read the opaque written language of English. Phonological awareness was related to word decoding in English for both first and second grade. In contrast, phonological awareness was related to word decoding in Greek for first grade but no longer for second grade. What appears to be the case is that phonological awareness is important at the beginning phase of learning to read regardless of the nature of the alphabetic script. However, the beginning phase of learning to read is transited more quickly by learners of transparent scripts than by learners of opaque ones.

CONCLUSIONS

Our understanding of dyslexia has evolved from the early myths about backward reading to the current view that dyslexia involves a problem in word-level reading resulting from issues in the processing of language. This evolution illustrates several ideas with implications that extend well beyond the field of dyslexia.

First, there are multiple potential sources of developmental and individual differences for complex cognitive tasks. Normal reading requires adequate vision: eye-movement control; visual perception; knowledge of the orthographic, phonological, and semantic features of words; higher-level linguistic and cognitive skills; memory; and attention. If any of these multiple potential determinants of performance is seriously affected, reading performance will suffer. For example, normal reading is not possible with profound visual impairment unless an alternative system such as Braille is made available, and children with serious problems in attention typically are impaired in comprehending long and involved passages. Nevertheless, the results of studies of the normal development of reading and of the atypical development of dyslexia indicate that deficits in language rather than in vision constrain performance. The language problem appears to be identifiable by poor performance on phonological awareness tasks. The good news for educators and others involved in helping individuals with dyslexia is that most individuals respond to phonological awareness interventions with improvements in both phonological awareness and reading.

It certainly is the case that individual differences exist in essential components of normal reading, such as visual acuity. But except for cases of profound visual impairment, it is a myth that individual differences in visual acuity affect reading performance, because the acuity requirements of normal reading are quite modest relative to human capabilities in visual activity. We can of course construct a reading task for which acuity is a determining factor. The vision charts that optometrists used to ask us to read on a distant wall, and which now are simulated in a visual display, provide an obvious example. But when we can adjust reading material so that it is at a comfortable distance, acuity largely becomes irrelevant.

The second idea is that careful experiments involving comparisons of normally developing and affected individuals is key to isolating critical determinants of performance. For the case of dyslexia, the observation that normally developing beginning readers made as many reversal errors as did older readers with dyslexia was crucial for recognizing the myth that dyslexia was caused by seeing things reversed. Comparing the eye movements of normal readers when given difficult reading material to the eye movements characteristic of poor readers had a similar result in refuting the myth that dyslexia results from faulty eye movements. A number of other cases could have been discussed. For example, when one of us was trained as a school psychologist nearly 30 years ago, scatter in the profile of scores on an IQ test was considered to be a diagnostic sign of a specific learning disability. This seemed logical at the time because, by definition, a learning disability was considered to be a specific deficit among other areas of strength. Studies of individuals with learning disabilities and the personal experience of practitioners provided empirical support to this view: Considerable variability actually existed in the profile of scores provided by individuals with learning disabilities. The crucial piece of disconfirming evidence for this myth was provided by an analysis of how much scatter was typical in the standardization sample of a popular IQ test. It turned out that scatter was the norm rather than the exception.

We need to acknowledge that our analysis of the problem of dyslexia was necessarily limited in its scope. Other important determinants of reading performance exist. Perhaps the most important of these is the probability that reading problems are exacerbated by reading instruction that is largely directed toward the typical student and becomes increasingly ineffective for children who are behind (Spear-Swerling & Sternberg, 1996).

In closing, the moral of the story of deciphering dyslexia is well captured by the healthy skepticism advocated by a former president of the United States:

> The greatest enemy of the truth is not the lie—deliberate, contrived, and dishonest, but the myth—persistent, pervasive, and unrealistic.
>
> —John F. Kennedy

REFERENCES

Aaron, P. G. (1989). Qualitative and quantitative differences among dyslexic, normal, and nondyslexic poor readers. *Reading & Writing, 1,* 291–308.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14,* 255–259.

Anthony, J. L., Lonigan, C. J., Burgess, S. R., Driscoll, K., & Phillips, B. M., (2002). Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology, 82,* 65–92.

Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M., & Burgess, S. R. (2003). Preschool phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly, 38,* 470–487.

Baddeley, A. (1982). Reading and working memory. *Bulletin of the British Psychological Society, 35,* 414–417.

Baddeley, A. (1986). *Working memory.* Oxford: Oxford University Press.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26,* 49–66.

Bloomfield, L. (1933). *Language.* Oxford, England: Holt.

Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia, 44,* 26–59.

Bus, A. G., & van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology. 91,* 403–414.

Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology, 81,* 313–321.

Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology, 83,* 451–455.

Byrne, B., & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology, 85*(1), 104–111.

Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology, 87,* 488–503.

Byrne, B., Fielding-Barnsley, R., & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology, 92,* 659–667.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91,* 77–111.

Chall, J. S. (1967/1983). *Learning to read: The great debate.* New York: McGraw-Hill.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55,* 75–84.

Crowder, R. G., & Wagner, R. K. (1992). *The psychology of reading: An introduction* (2nd ed.). London: Oxford University Press.

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the national reading panel's meta-analysis. *Review of Educational Research, 71,* 393–447.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., 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.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities.* New York: Guilford.

Foorman, B. R., Chen, D.-T., Carlson, C., Moats, L., Francis, D. J., & Fletcher, J. M. (2003). The necessity of the alphabetic principle to phonemic awareness instruction. *Reading and Writing: An Interdisciplinary Journal, 16,* 289–324.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90,* 37–55.

Foorman, B. R., Francis, D. J., Novy, D. M., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology, 83,* 456–469.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal individual growth curve analysis. *Journal of Educational Psychology, 88,* 3–17.

Gelb, I. J. (1952). *A study of writing.* Chicago: University of Chicago Press.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology, 100,* 566–580.

Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia, 52,* 141–163.

Goulandris, N. (Ed.). (2003). *Dyslexia in different languages: cross-linguistic comparisons.* London: Whurr.

Hallahan, D. P., & Mock, D. R. (2003). A brief history of the field of learning disabilities. In H. L. Swanson & K. R. Harris (Eds.), *Handbook of learning disabilities* (pp. 16–29). New York: Guilford.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65,* 41–57.

Hulme, C., Snowling, M., Caravolas, M., & Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9,* 351–365.

Jorm, A. F., & Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics, 4,* 103–147.

Lonigan, C. J., (2006). Conceptualizing phonological processing skills in prereaders. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early childhood literacy research* (2nd ed., pp. 77–89). New York: Guilford.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *Test of preschool early literacy.* Austin, TX: Pro-Ed.

Lovett, M. W., Steinbach, K. A., & Frijters, J. C. (2000). Remediating the core deficits of developmental reading disability: A double-deficit perspective. *Journal of Learning Disabilities, 33,* 334–358.

National Center for Educational Statistics (2006). *Digest of education statistics, 2005* (NCES 2006–030), Chapter 2. Washington, DC: U.S. Department of Education.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the national early literacy panel*. Washington, DC: National Institute for Literacy.

National Institute of Child Health and Human Development. (2000). *Report of the national reading panel: Teaching children to read.* Washington, DC: U.S. Department of Health and Human Services.

National Research Council. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly, 24,* 119–134.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2,* 31–74.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Upper Saddle River, NJ: Prentice Hall.

Schneider, W., Ennemoser, M., Roth, E., & Kuespert, P. (1999). Kindergarten prevention of dyslexia: Does training in phonological awareness work for everybody? *Journal of Learning Disabilities, 32,* 429–436.

Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology, 92,* 284–295.

Seki, A., Kassai, K., Uchiyama, H., & Koeda, T. (2008). Reading ability and phonological awareness in Japanese children with dyslexia. *Brain and Development, 30,* 179–188.

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology, 59,* 451–475.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in your children.* Washington, DC: National Academy Press.

Spear-Swerling, L., & Sternberg, R. (1996). *Off track: When poor readers become learning disabled,* Boulder, CO: Westview.

Stanovich, K. E. (1982). Individual differences in the cognitive processes of reading: I. Word decoding. *Journal of Learning Disabilities, 15,* 485–493.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disability: A regression-based test of the phonological core variable-difference model. *Journal of Educational Psychology, 86,* 24–53.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34,* 33–58.

Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 161–188). Mahwah, NJ: Erlbaum.

Torgesen, J. K., Rashotte, C. A., & Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolr (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–355). Parkton, MD: York.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91,* 579–593.

Tzeng, S.-J. (2007). Learning disabilities in Taiwan: A case of cultural constraints on the education of students with disabilities. *Learning Disabilities Research & Practice, 22,* 170–175.

Vellutino, F. R. (1979). *Dyslexia: Theory and research.* Cambridge, MA: MIT Press.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of School Psychology, 36,* 367–397.

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.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processes.* Austin, TX: PRO-ED.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33,* 468–479.

Werker, J. F., Bryson, S. E., & Wassenberg, K. (1989). Toward understanding the problem in severely disabled readers: II. Consonant errors. *Applied Psycholinguistics, 10,* 13–30.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848–872.

Wise, B. W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology, 72,* 271–304.

Wydell, T. N. (2003). Dyslexia in Japanese and the hypothesis of granularity and transparency. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 255–276). London: Whurr.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3–29.

PART II

Mind Meets the Classroom

This page intentionally left blank

3 Insights From Everyday Knowledge for Mathematics Education

TEREZINHA NUÑES AND PETER BRYANT

"For a mathematician, a proportion is a statement of equality of two ratios, i.e., $a/b = c/d$" (Tourniaire & Pulos, 1985, p. 181). This mathematical definition should not blind anyone, and in particular educators, to proportional knowledge that does not come dressed in this algebraic robe. The aim of this chapter is to discuss children's and adults' understanding of proportions when their knowledge is not expressed in algebraic form. We seek here to identify the origins and characteristics of this knowledge and discuss its potential and limitations should it be taken up in school.

This chapter follows the path of our own discoveries about proportional reasoning. Our interest on proportional reasoning started from research in developmental psychology. Proportional reasoning was initially investigated by Piaget and his colleagues (Inhelder & Piaget, 1959; Piaget & Inhelder, 1975), who described the schema of proportionality as an achievement of the period of formal operations. Our question was whether this schema could be constructed by people with little or no schooling. The search for proportional reasoning outside school is the focus of the first section in this chapter. Once we found that proportional reasoning can be achieved without schooling, we were then puzzled by a question: Why is it that unschooled people can construct this schema in their everyday occupations yet students with all the benefit of mathematics instruction find this concept so difficult? The second section of this chapter reviews research by colleagues in mathematics education, which describes students' solutions to proportions problems in school settings. The third and final section is to a large extent theoretical: it examines the obstacles to the use of students' informal reasoning in school and theories that might help us find ways to remove these obstacles.

PROPORTIONAL REASONING OUTSIDE OF SCHOOL

The diagram presented in Figure 3.1 illustrates the simplest type of proportions problems, which most people come across in everyday life. In simple proportions problems there are two quantities—in this case, the number of pineapples and the price—and there is a fixed ratio between the two, indicated by the arrows in the diagram. The problem in Figure 3.1 is to calculate how much one must pay for 12 pineapples.

In proportions problems in school, students are not usually given the unit value—that is, the price of one pineapple; when the unit price is given, the problem is called a "multiplication problem." However, from the viewpoint of the mathematical relations in the situation, there is no distinction between multiplication and proportions problems, because both involve the assumption of a fixed ratio between the two quantities.

Piaget and his colleagues (Inhelder & Piaget, 1959; Piaget & Inhelder, 1975) did not investigate children's understanding of proportionality in situations like the one in Figure 3.1. The problems they

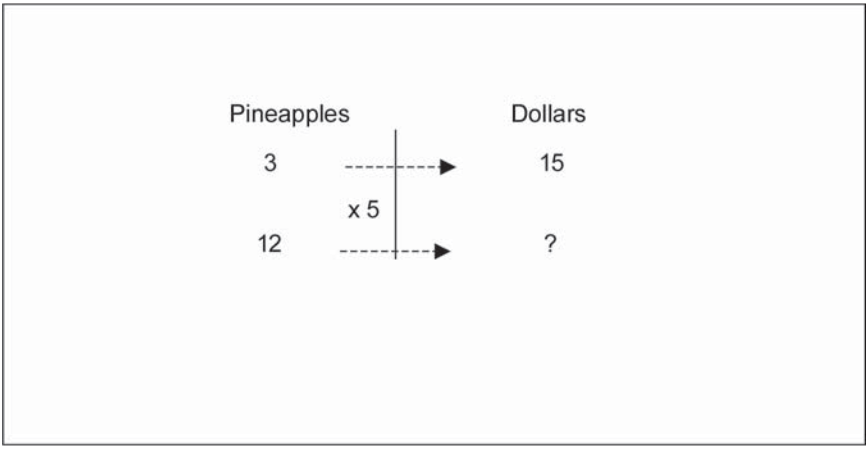


Figure 3.1 A schematic representation of the simplest type of proportions problems.

presented to the children and adolescents in their studies were much more complex: they involved multiple proportions situations, in which one variable (or quantity, as described in Figure 3.1) is proportional to two or more other variables. These problems also involved direct as well as inverse proportions. In one study, for example, the children were asked to make predictions about the size of the shadow projected on a screen by an object placed between a light and the screen. The size of the shadow is directly proportional to the size of the object (the bigger the object, the larger the shadow) and inversely proportional to the distance it is placed from the light (the closer to the light, the larger the shadow). In another study, the children were asked to compare the probability of drawing a card marked with a cross from two groups of cards with different numbers of marked and unmarked cards. The probability of success—that is, of drawing a marked card—is directly proportional to the number of marked cards and inversely proportional to the number of unmarked cards. In a third study, they asked children to balance weights on a T-shaped scale, but putting different numbers of weights on each side of the scale. In this problem, the scale will balance if the number of weights multiplied by their distance from the fulcrum is the same. So 2 pieces placed in the third position from the fulcrum will balance with 3 pieces placed in the second position. Two variables, number of weights and position on the scale's arm, must be considered by the children in order for them to succeed in the task without using trial-and-error procedures.

Piaget and his colleagues did not ask the children to solve problems presented with numbers only. Their problems involved judgments (as in the probabilities task) or actions (as in balancing weights on a scale). In this sense, these problems were rather different from the types of proportions problems that students are asked to solve in school. Because the problems were presented with materials that the children and adolescents could manipulate as they tried to find solutions, they allowed the children to draw on knowledge that they might not have been able to explain in words or represent with mathematical symbols. For example, in the probabilities task, the children were presented with an item where two groups of cards were compared: in group 1, there was 1 card marked with a cross and 2 unmarked cards; in group 2, there were 2 marked and 4 unmarked cards. Some children succeeded in understanding the equivalence of the probabilities by manipulating the cards and showing that in both groups there were 2 unmarked cards for each card with the cross. They did not have to say that 1/3 = 2/6 in words or using numerical representations: they were able to use a "theorem in action" (see Vergnaud, in press) that could be put into words as: "if there is a 1 to 2 ratio of marked to unmarked cards, the total number of cards does not matter and the probability of drawing a marked card is the same." This theorem provides the basis for understanding the mathematical model in proportions problems.

Piaget's studies indicated that young adolescents, at about age 11 to 15, were able to understand the sorts of proportions problems that Piaget and his colleagues had devised. They distinguished two steps that the children took in this process. In the first step, the children established ratios between the variables by manipulating the materials, as in the example in the previous paragraph regarding probabilities. This was termed an *empirical approach* to solution. In the second step, the young adolescents were able to express the equivalence using symbolic representations, either verbally or numerically. However, Piaget did not view this as a universal result in cognitive development. In a classic paper, Piaget (2008; first published in 1972) argued that his studies formed the basis for describing "the logic of cultured adults" and that

> the rate at which a child progresses through developmental succession may vary, especially from one culture to another. Different children also vary in terms of the areas of functioning to which they apply formal operations, according to their aptitudes and their professional specializations. (p. 40)

The ages at which students seemed to conquer Piaget's proportionality tasks were very similar to those at which, in many countries, they receive explicit teaching about proportions. Consequently, and inspired by Piaget's own comments, we became interested in finding out to what extent mathematics instruction about proportions explained the emergence of proportional reasoning and was thus a product of schooling rather than a universal phase in cognitive development.

We reasoned that we needed to find professional contexts in which adults who had not attended school long enough to be taught about proportions were nevertheless, by the nature of their occupations, required to reason proportionally. We also thought that multiple proportions problems, where two variables are linearly related to a third one, might be unnecessarily complex: simple proportions problems, where only one variable affects the target one, might demonstrate that the schema of proportionality develops without the benefit of school instruction.

In a series of studies (summarized in Nunes, Schliemann, & Carraher, 1993), we asked Brazilian students who had received instruction about proportions as well as adults who had not attended school long enough to receive such instruction (and sometimes had not attended school at all) to solve simple proportions problems. These studies involved participants in a large city (Recife, in the northeast of Brazil), which had 2.5 million people, and in a small town (Itapissuma), which had only one secondary school. The students in these studies were in their seventh year of school. The adults were working in contexts rather different from each other: one group was composed of foremen at construction sites and a second of fishermen. They had little or no schooling owing to their social status. The questions we asked them were about proportions in situations related to their work.

Foremen in the construction industry have to work with blueprints as representations of distances in the buildings under construction. They have experience with a certain number of conventionally used scales (e.g., 1:50, 1:100, and 1:1000). We gave them scale drawings that involved different ratios (e.g., 1:40), and the ratio used in the drawing was not indicated; the information they could extract from the blueprint was that a wall that measured 5 cm in the blueprint had its real size indicated as 2 m. Most foremen were able to use this information and to calculate the measure of another wall from its measure on the blueprint (for further details, see Nunes Carraher, 1986). Completely illiterate foremen ($n = 4$), who had never set foot in a school owing to their life circumstances, showed 75% correct responses to this type of problem. In contrast, the seventh-grade students, who had been taught to solve proportions problems using the formal method known as the rule of three, which involves writing an equation of the form $a/b = c/d$ and solving for the unknown value, performed significantly worse (60% correct).

The results for the Brazilian students do not differ much from those observed in other parts of the world: students have modest rates of success in solving proportions problems. The results of the Third International Mathematics and Science Survey show, for example, that the rates of success for students in their eighth year in school are quite similar: Danish students succeeded on average in 47% of the proportionality problems and Swedish students showed a 50% success rate; in the German speaking part of Switzerland, the rate of success was 60% (Beaton, Mullis, Martin, Gonzalez, Kelly, & Smith, 1996). However, the results of our study were surprising. The foremen, who did not participate in "the cultured adults" world of Brazilian society, were more successful in solving problems for which they did not have a routine than were students who had been taught a general procedure to solve proportions problems.

Nunes, et al. (1993) showed that fishermen and foremen who had little formal school instruction could solve proportions problems that were novel to them in three ways: (a) the problems used values that departed from the values they normally worked with (e.g., the 1:40 scale referred to earlier); (b) they were asked to calculate in a direction which they normally did not have to think about (e.g., fishermen are used to calculating the amount of shelled shrimp which you get from specific quantities of unprocessed shrimp but not how much they needed to catch if a customer wanted a specific amount of processed seafood); or (c) the content of the problem was different from the problems with which they worked in their everyday lives (e.g., fishermen were asked to calculate how much flour you could get from specific amounts of cassava).

A salient feature of the method used both by fishermen and foremen in these calculations was the way in which they consistently established relations between the two quantities in the problem and explicitly referred to these quantities. They expressed these correspondences between the quantities orally, with no objects or written records to support their reasoning. For example, the fishermen were asked: "There is a type of shrimp in the South that yields 3 kilos of shelled shrimp for every 18 kilos that you catch; how many kilos would you have to catch for a customer who wants 2 kilos of shelled shrimp?" An example of the responses offered by the fishermen is presented in Figure 3.2, which includes his oral solution at the left and a diagram of the relations between the quantities in the right column.

This form of reasoning makes explicit references to the quantities: there is no reference to the ratio of shelled to unprocessed shrimp—that is, there is no reference to the fact that the amount of unprocessed shrimp is 6 times the amount of shelled shrimp.

One could argue that perhaps this fisherman knew the 2- and 3-times tables but not the 6-times table. This is unlikely, as the calculations that he used—one-half of 3 is 1.5 and one-third of 1.5 is one-half—are not part of the times tables. However, we did check the possibility that the difficulty of the numerical values was important by giving the fishermen problems that were carefully designed to test this hypothesis. Figure 3.3 presents a diagram of how this design was implemented across problems: the numerical relations within one quantity in one problem are swapped around in the other problem, to represent the ratio between

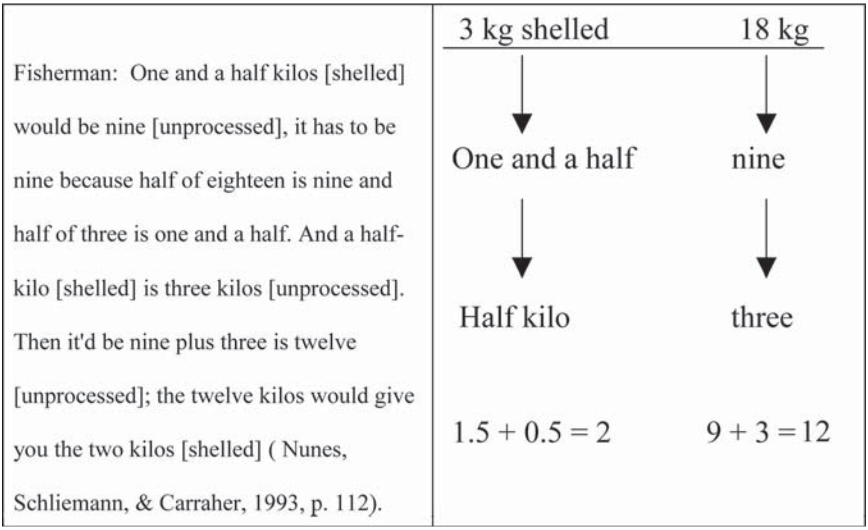


Figure 3.2 A fisherman's solution to a proportions problem.

the quantities. In order to test whether the fishermen found it easier to think about quantities than about ratios, we used the same numbers in different arrangements across the different problems.

To our surprise, although they were more successful with problems of type A than type B, this difference was not statistically significant: their rate of correct responses to problems of type A was 80% and to problems of type B, 75%. However, they were quicker when they solved the type A problems, as the quantities were multiples or divisors of each other, and this meant that they used fewer steps in calculation. In contrast, the students' performance surprised us. Although they had learned, in school, the algebraic procedure to solve any proportions problem, they were significantly more successful in solving type A than type B problems: their performance dropped from 80% correct in type A to 35% correct in type B problems.

In brief, our research about the schema of proportionality started with the question raised by Piaget about the possible lack of generality of his findings because he had studied participants who shared the logic of cultured adults. What we found was that exposure to this logic in the algebraic format $a/b = c/d$ did not give students an advantage over adults who had developed their understanding of proportionality outside school.

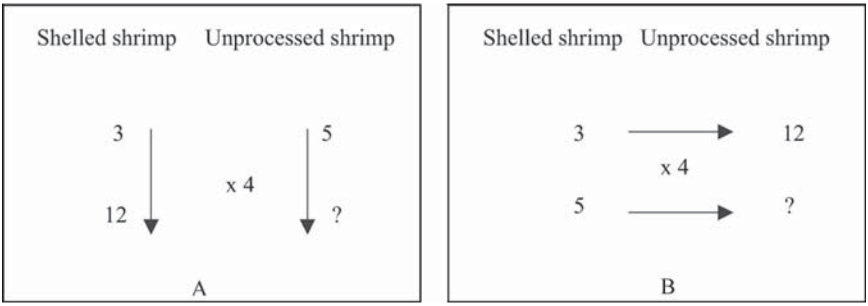


Figure 3.3 Two proportions problems with the same numbers but different levels of difficulty.

In order to understand students' performance, we now turn to results of research with students in proportions problems in countries where this topic is taught in school from the age of about 10 years.

PROPORTIONAL REASONING AMONG STUDENTS

Piaget's studies of the development of the schema of proportionality led to a large number of investigations by mathematics educators (e.g., Hart, 1981, 1984; Karplus & Peterson, 1970; Karplus, Pulos, & Stage, 1983; Lamon, 1994; Lesh, Post, & Behr, 1988; Noelting, 1980a, 1980b; Ricco, 1982; Vergnaud, 1983). Perhaps these educators expected to find that Piaget was entirely wrong: after all, why should students have difficulty with proportions problems after they had been taught simple routines to solve these problems? In particular, it should be noted that the arithmetic in these problems should not challenge students: to solve for the unknown using the formula $a/b = c/d$, students only need to use multiplication and division, and they will have practiced procedures to calculate multiplication and division for about 3 to 4 years before being taught about proportions in school.

Whatever their expectations, researchers found that students' thinking was quite impervious to school instruction on how to solve proportions problems. Hart (1981), for example, presented the following problem to a large sample of students (2,257) aged 11 to 16 years in 1976: 14 meters of calico cost 63 p; find the price of 24 meters. She reported that no student actually quoted a method that he or she had been taught in school to solve this problem, which is to find the unitary ratio by dividing 63 by

14 and then multiplying the result (4.5) by 24. In France, where students are taught the rule of three (also taught in Brazil and described above), most students do not use it to solve proportions problems (Vergnaud, 1983). Hart (1981) also reports that the formula $a/b = c/d$ was taught to 100 students in one school where she carried out her investigations of proportions problems, but it was used by only 20 students, 15 of whom were among the high achievers in the school.

So the assimilation of taught procedures is modest. What do students do instead? Many researchers (e.g., Hart, 1981; Kaput & West, 1994; Lamon, 1994; Nunes & Bryant, 1996; Ricco, 1982; Steffe, 1994) have described students' solutions to proportions problems, and different terms have been used to refer to the solution procedures: building-up strategies, empirical strategies, halving or doubling, and replications of a composite unit. In essence, these solutions procedures consist in using the initial values provided in the problem and changing them in one or more steps to arrive at the desired value. Hart's (1981) well-known example of the onion soup recipe for 8 people, which has to be converted into a recipe for 6 people, illustrates this strategy well: Students find what the ingredients are for 4 people, which is half of what is needed for 8 people; 4 people plus half of 4 makes 6 people, so the children take each of the ingredients in the recipe in turn, halve the amount, and add this to the amount required for 4 people. The similarity between English students' solutions and those offered by Brazilian fishermen and foremen cannot go unnoticed.

These solution procedures are known as informal strategies—that is, they are not taught in school, so they must have been acquired outside of school. Their ubiquitous presence should provoke reflection. We discuss two aspects of this informal knowledge: the first is its origin and the second is why it goes unnoticed in school until teachers start to teach students about proportion.

The origin of these informal strategies is most likely to lie in the schema of one-to-many correspondences. Piaget (1952) seems to be the first researcher to have carried out research on children's use of one-to-many correspondences to solve problems. But since then many other investigations have been carried out. We describe here two studies in which children who had not yet been taught about multiplication and division in school were asked to solve multiplicative reasoning problems—that is, problems in which there is a fixed ratio between two quantities, which in school would be identified as multiplication or division problems.

Kouba (1989) presented young children in the United States in first, second, or third grade (about 6 to 8 years of age), with multiplicative reasoning problems typical of those used later in school; for example: at a party, there were 6 cups and 5 marshmallows in each cup; how many marshmallows were there? The key characteristic of proportions situations is undoubtedly present in this problem: a fixed ratio between cups and marshmallows.

Kouba (1989) analyzed the children's strategies in great detail and classified them in terms of the types of actions used and the level of abstraction. The level of abstraction varied from direct representation (i.e., all the information was represented by the children with concrete materials), through partial representation (i.e., numbers replaced concrete representations for the elements in a group and the child counted in groups), up to the most abstract form of representation available to these children (i.e., the use of numbers only during the solution process).

For the children in first and second grade, who had not received instruction on multiplication and division, the most important factor in predicting their solutions was which quantity was unknown. For example, in the problem above, about the 6 cups with 5 marshmallows in each cup, when the size of the groups was known (i.e., the number of marshmallows in each cup), the children used correspondence strategies: they paired objects (or tallies to represent the objects) and counted or added, creating one-to-many correspondences between the cups and the marshmallows. For example, if they needed to find the total number of marshmallows, they pointed 5 times to a cup (or its representation) and counted to 5, paused, and then counted from 6 to 10 as they pointed to the second cup, until they reached the solution.

In contrast, when the number of elements in each group was not known, the children used dealing strategies: they shared out one marshmallow (or its representation) to each cup, and then another, until they reached the end, and then counted the number in each cup.

Although the actions just described look quite different, their aims are the same: to establish one-to-many correspondences between the marshmallows and the cups.

Kouba observed that 43% of the strategies used by the children, including first, second, and third graders, were appropriate. Among the first- and second-grade children, the overwhelming majority of the appropriate strategies was based on correspondences, either using direct representation or partial representation (i.e., tallies for one variable and counting or adding for the other); few used recall of multiplication facts.

The recall of number facts was significantly higher after the children had received instruction in third grade.

The level of success observed by Kouba among children who had not yet received instruction is modest compared to that observed in two subsequent studies, where the ratios were easier. Becker (1993) asked kindergarten children in the United States, aged 4 to 5 years, to solve problems in which the correspondences were 2:1 or 3:1. The children were more successful with 2:1 than 3:1 correspondences, and the level of success improved with age. The overall level of correct responses by the 5-year-olds was 81%. Carpenter, Ansell, Franke, Fennema, and Weisbeck (1993) also gave multiplicative reasoning problems to U.S. kindergarten children involving correspondences of 2:1, 3:1, and 4:1. They observed 71% correct responses to these problems.

These success rates leave no doubt that many young children start school with some understanding of one-to-many correspondence, which they can use to learn to solve multiplicative reasoning problems in school. This is likely to be the basis for the building-up strategies observed by many different researchers in different countries.

These results do not imply that children who use one-to-many correspondence to solve multiplicative reasoning problems consciously recognize that in a multiplicative situation there is a fixed ratio connecting the two variables. Their actions maintain the ratio fixed, but it is most likely that this invariance remains, in Vergnaud's (1997) terminology, a "theorem in action." The children's attention is focused on the quantities, which they add, double, or halve according to the circumstances and the data in the problem.

The second issue we raised was the invisibility of children's informal knowledge of multiplication in school. We have only a speculative answer to why children's use of one-to-many correspondences goes unnoticed until they start to use these strategies to solve proportions problems. We speculate that the way in which children are taught multiplication in school might explain why building-up or correspondence strategies are described only in the context of students' solutions to proportions problems.

In school, in the United Kingdom at any rate, students are taught that multiplication is repeated addition. The focus of this teaching is on calculation: because one can use repeated addition to calculate a product, this view of multiplication works within the context of calculation. However, different researchers (e.g., Confrey, 1994; Kieren, 1994; Nunes & Bryant, 1996; Vergnaud, 1983) have argued that additive and multiplicative reasoning are distinct and that it is important for students to understand this.

In additive reasoning situations (i.e., situations where addition or subtraction are used to solve problems), the relevant schemas of action are joining and separating quantities. For example, you have 3 red balloons and 5 blue balloons; how many do you have all together? Or you have 8 balloons and give 3 away; how many do you have left? In multiplicative reasoning, the relevant schema required to understand the problem is setting quantities in correspondence, with a fixed ratio between them: 1 pineapple costs $5. We suggest that if teachers and children alike focus their attention on a single quantity, the $5 in this example, and ignore the fixed ratio between dollars and pineapples, the schema of correspondence goes unnoticed because its main characteristic is that it keeps both variables in focus.

However, we believe that this is not the whole story: the difficulty of multiplicative reasoning and mastering proportionality is not simply engendered by teaching. There is a huge conceptual step to be taken by students in order to master proportional (or, more generally, functional) reasoning. This step is explored in the next section of this chapter.

NUMBERS, QUANTITIES, AND RELATIONS

Numbers are symbols that we use to represent (among other things) quantities and relations. Thompson (1993) has suggested that

> a person constitutes a quantity by conceiving of a quality of an object in such a way that he or she understands the possibility of measuring it. Quantities, when measured, have numerical value, but we need not measure them or know their measures to reason about them. You can think of your height, another person's height, and the amount by which one of you is taller than the other without having to know the actual values. (pp. 165–166)

Children experience and learn about quantities and the relations between them quite independently of learning to count. Similarly, they can learn to count quite independently from understanding quantities and relations between them. A significant insight that they need to achieve in the development of their mathematical reasoning is that numbers are used to represent quantities. Thus, if two quantities are equivalent, they should be represented by the same number; conversely, if two quantities are represented by the same number, they are equivalent.

Piaget's (1952) studies of children's understanding of number showed quite clearly that some children who know that two sets are equivalent do not deduce the number in one set if they have counted the other. More recent research offers quantitative information on 4-year-olds' ability to infer how many sweets are in a set after having counted the sweets in an equivalent set. Frydman and Bryant (1988) asked 4-year-old children to share a set of "chocolates" between two recipients. At this age, children often share things between themselves, and they typically do so on a one-for-A, one-for-B, one-for-A, one-for-B basis. When the child had done the sharing, the experimenters counted out the number of items that had been given to one recipient, which was 6. Having done this, they asked the child how many chocolates had been given to the other recipient. None of the children immediately made the inference that there were 6 chocolates in the second set, even though they had just stated that the sharing had been fair, and both recipients had the same amount of chocolate. Instead, every single child began to count the second set. In each case, the experimenter then interrupted the child's counting and asked him or her if there was any other way of working out the number of items in the second recipient's share. Only 40% of the 4-year-olds made the correct inference that the second recipient had also been given 6 chocolates.

Sarnecka and Gelman (2004) recently replicated this observation using a different method. They report that 3- and 4-year-olds know that if a set had 5 objects and you add some to it, it no longer has 5 objects; however they did not know that equal sets must have the same number word.

These results show that children do not necessarily know that quantities that are equivalent are represented by the same number. An important step that children must take in the use of numbers to represent quantities is the understanding that quantities that are equivalent are represented by the same number and, conversely, those that are represented by the same number are equivalent and need not be compared directly. Although it takes children some time to establish this connection between numbers and quantities, most seem to succeed in doing so in the early years in primary school.

There is no doubt that children must grasp how numbers and quantities are connected in order to understand what numbers mean. But numbers are not used in mathematics only to represent quantities.

A major use of mathematics is to manipulate numbers that represent relations and arrive at conclusions without having to operate directly on the quantities. Attributing a number to a quantity is measuring; quantifying relations and manipulating them is quantitative reasoning or modeling. Thompson (1993) defined "Quantitative reasoning [as] the analysis of a situation into a quantitative structure—a network of quantities and quantitative relationships" (p. 165). Elsewhere, Thompson (1994) emphasized that "a quantitative operation is nonnumerical; it has to do with the *comprehension* of a situation" (pp. 187–188).

In the previous sections of this chapter we described fishermen's, foremen's, and students' informal knowledge of proportions. We argued that, as they spoke about their solutions, they spoke about quantities, and not about the fixed ratio between them. This analysis led us to hypothesize that they used a form of reasoning that focuses on quantities; their knowledge of the fixed ratio between the quantities remains implicit. The ratio is conserved constant during their solution procedures because of the manipulations that they carry out on the quantities: if two quantities are doubled (or halved, or multiplied by 3), the ratio between them remains the same. The fishermen and foremen in our studies were successful in problems where thinking of the ratio offered an easy road to solution—not because they thought of the ratio but because they were apt in computations and used longer and awkward routes to solve the problems when a single simple calculation would have led to the answer. Like the fishermen and foremen, students who solved Hart's onion soup problem by a building up strategy (adapting the recipe for 4 to 6 people by halving the ingredients and then adding the ingredients for 4 to those required for 2 people), were most likely thinking of the quantities of flour, stock cubes, water, and so on. The idea of a fixed ratio between quantities remains implicit when the focus is on the quantities.

Fischbein (1987) argued that the difficulty of implicit models is that they can be overgeneralized exactly because they are implicit; a conscious choice of a model for relations in a problem can be based only on knowledge that is to some extent explicit. Thus students would be expected to generalize the use of solution procedures used for proportional problems to other problems where the relation to be considered is a different one. There is indeed some evidence that students use additive reasoning in multiplicative problems as well as multiplicative reasoning in additive problems.

Many researchers have reported that students use additive reasoning where they should have used multiplicative reasoning. Karplus, et al.

(1983), for example, presented students with a problem where two figures, Mr. Short and Mr. Tall, were measured with buttons; Mr. Short was also measured with paper clips. Students were asked what Mr. Tall's height in paper clips would be. This is a scale problem, where the scale was 1:1.5 (one paper clip was equivalent to 1.5 buttons), but many students calculated the difference in height when the two figures were measured with buttons and added this difference to Mr. Short's measure in order to determine Mr. Tall's height in paper clips.

The use of multiplicative reasoning when an additive relation was the relevant one has also been documented. Cramer, Post, and Currier (1993) report that prospective teachers in the United States used multiplicative reasoning to solve this problem: Two girls, Sue and Julie, are running on a track at the same speed. Sue started first. When she had run 9 laps, Julie had run 3 laps. When Julie completed 15 laps, how many laps had Sue run? The majority of the prospective teachers wrongly answered 45 laps, taking the ratio between 3 and 9 as the relevant relation to the solution of this problem. There is, in fact, a fixed ratio in the problem, but it is 1:1, because Sue and Julie were running at the same speed, so when Julie runs one lap, Sue also runs one lap. The relation that is relevant to the solution of this problem is the difference between the two: Sue started first and was 6 laps ahead of Julie. It is possible that the prospective teachers overgeneralized the use of a method for solving proportions problems because they had not developed sufficient awareness of the significance of fixed ratios in such situations.

THE ROLE OF INFORMAL KNOWLEDGE IN LEARNING MATHEMATICS IN SCHOOL

Some researchers, such as Fischbein (1987), Streefland (1984), and Treffers (1987), maintain that children's informal knowledge is a starting point for learning mathematics in school. This is assumed to be the case in a large variety of situations, including in learning about proportions. Others, such as Booth (1981) and Adjiage and Pluvinage (2007), suggest that implicit models are an obstacle to learning. Finally, Resnick (1983) and Kaput and West (1994) do not explicitly take either side for the argument but note that an important lesson from psychological and mathematics education research is that, even after people have been taught new concepts and ideas, they still resort to their prior methods to solve problems that differ from the textbook examples on which they have applied their new knowledge. Irrespective of whether they help or hinder learning, these methods must not be ignored in the classroom.

A precise answer to the question of whether informal knowledge hinders or helps learning may depend on what type of teaching children receive in school. However, it is possible to obtain relevant evidence by means of a longitudinal study. If informal knowledge does in fact offer a starting point for learning mathematics in school, children who start school with more informal knowledge of multiplicative reasoning should have an advantage in learning mathematics.

We (Nunes et al., 2007) carried out a longitudinal study that addressed this question. We wanted to know whether the children's ability to use one-to-many correspondences to solve multiplicative reasoning problems at the start of school gave them an advantage in learning mathematics. Multiplicative reasoning is relevant in many aspects of mathematics learning, not just solving multiplication, division, and proportions problems. Multiplicative reasoning is, for example, implicit in understanding place value (Kamii, 1981): when we write the number 56, the digit 5 represents 5×10, a multiplication that the children may understand implicitly or explicitly.

The children were tested on their understanding of multiplicative reasoning at the start of school, when they were between 5 and 6 years old. They were asked to solve, among others, five multiplicative reasoning problems that could be solved by one-to-many correspondence. About 14 months later, the children were given a state-designed, teacher-administered mathematics achievement test, which is entirely independent of the researchers and an ecologically valid measure of how much they have learned in school.

It is quite possible that young children who have developed this informal knowledge of multiplicative reasoning before starting school are just cleverer than other children or perhaps just better able to process information in problem-solving situations because of a better working memory, or, alternatively, they may just have more knowledge of numbers when they start school. In order to control for these alternative explanations, the children were given the British Abilities Scale (BAS-II; Elliott, Smith, McCulloch, 1997) to assess their general cognitive ability, a Working Memory Test, Counting Recall (Pickering & Gathercole, 2001), and, separately, a subtest of the BAS, Number Skills to assess their knowledge of numbers when they started school. These measures were then used in a fixed-order multiple regression analysis, which allowed us to control for the children's cognitive ability, working memory, and number knowledge and to test whether there is a specific relation between their informal knowledge of multiplicative reasoning at school start and their mathematics achievement a year later.

The results of this analysis are presented in the diagram in Figure 3.4, which displays the statistic (called Beta, with possible values between −1 and 1) that measures, in this mathematical model, the strength of the connection between each of the measures on the left and the children's mathematics achievement. The diagram shows clearly the importance of the children's knowledge of one-to-many correspondence, which is almost as high as their intelligence as measured by the BAS.

The total variance explained in the mathematics achievement by all the predictors was 66%; age explained 2% (nonsignificant), the BAS general score (excluding the Number Skills subtest) explained a further 49% ($p < .001$); the subtest on number skills explained a further 6% ($p < .05$); working memory explained a further 4% ($p < .05$), and the children's understanding of multiplicative reasoning at school entry explained a further 6% ($p = .005$). We therefore conclude that children's understanding

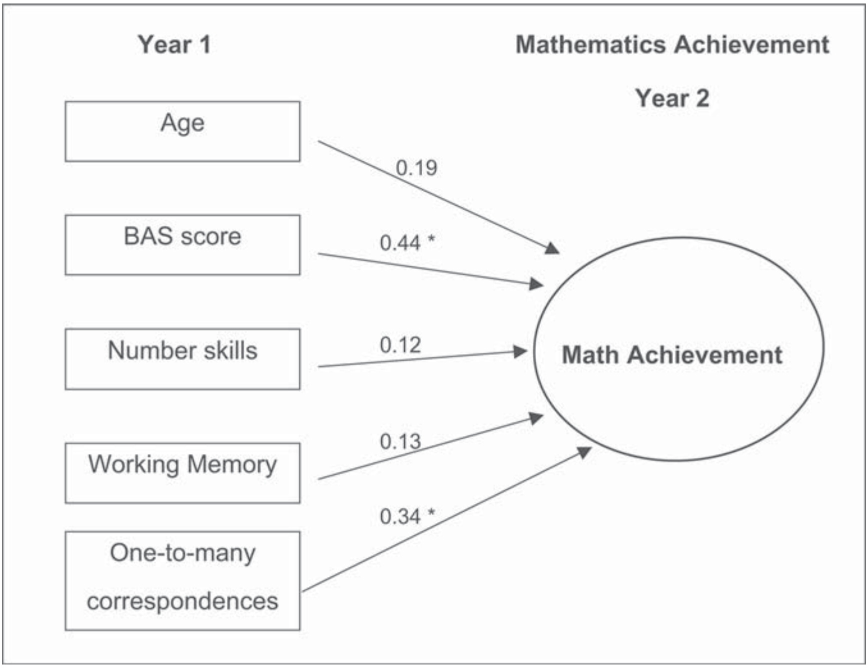


Figure 3.4 β values obtained in the model with all the predictions.

*A statistically significant value.

of multiplicative reasoning at school entry is a specific predictor of mathematics achievement in the first 2 years of school. These results support the view that, in a general way, this informal knowledge does offer a basis for learning mathematics in school.

We suggest that a more precise answer regarding the role of informal knowledge in school can be given only if different forms of teaching are compared. Freudenthal (1971) made this point forcefully by saying that if the mathematics teacher cuts the ties of mathematics to reality, those of us who do not become mathematicians will forget most of what we learned. In order for mathematics learning to have a lasting influence in our lives, it must be "tied to lived reality with strong bonds"—otherwise we return to the methods that are tied to our reality when we need to use mathematics to understand our world.

Fischbein (1987) has also argued that informal knowledge must be used as a basis for learning in the classroom. He suggested that the sort of informal and implicit knowledge of proportional relations that we have attributed to primary school students, which he called primary intuitions, plays an essential role in students' mathematical and scientific thinking processes. Primary intuitions seem to students self-evident and result in generalizations without any explicit analysis of what these implicit models mean. He adds that primary intuitions form the basis for secondary intuitions, which do not develop spontaneously "but through some educational intervention" (p. 71). They become intuitions later on, when students' understanding develops sufficiently for their new ideas to seem self-evident. If students were to explore the schema of one-to-many correspondence or their building up methods in the classroom, and if they were to become aware of the fact that they maintain a fixed ratio through their procedures (though they accomplish this implicitly), their informal knowledge could become a secondary intuition, where the ratio (or more generally the relation) between the quantities would play a major role.

The hypothesis that children reflect on their knowledge and then reconstruct it when they become aware of what previously they only knew implicitly is commonplace in psychological theories: it is part of general developmental theories, such as Piaget's theory of reflective abstraction (Piaget, 1978, 2001) and Karmiloff-Smith's theory of representational redescriptions in development (Karmiloff-Smith, 1992; Karmiloff-Smith & Inhelder, 1977). It is also used to describe development in specific domains such as language and literacy (Bryant & Bradley, 1985; Gombert, 1992; Karmiloff-Smith, 1992; Nunes & Bryant, 2009), memory (Flavell, 1971), and the understanding of others (Flavell, Green, & Flavell, 1990).

It is a reasonable hypothesis and it could also be pursued in mathematics education.

Different approaches have been suggested in mathematics education research regarding the best methods of making students aware of relations between variables. They include the use of graphic representations to scrutinize the nature of the relation visually (e.g., Ainley, Pratt, & Nardi, 2001), the use of tables to help students realize that the functional relation that connects the pairs of values in the two variables is a constant (e.g., Sellke, Behr, & Voelker, 1991), the use of metacognitive instruction (e.g., Kramarski, 2004), or promoting discussion in the classroom so that students verbalize their implicit models, becoming more aware of their own implicit ideas (e.g., Rojas-Drummond, Pérez, Vélez, Gómez, & Mendoza, 2003).

Kaput and West (1994) and also researchers in the Freudenthal Institute have over the years developed a program that includes all of these elements in the process that they call "the mathematization" of proportions situations. Some examples of the ideas put forward by the Freudenthal Institute researchers are briefly described.

Streefland and his colleagues (Streefland, 1984, 1985; van den Brink & Streefland, 1979) suggest that the first step in making students aware of relations between quantities might be taken through schematic drawing and visualization. Van den Brink and Streefland (1979) described a boy's reactions to proportions in drawings, where the intuition that something is not quite right can be taken up in conversation. The boy's reactions were taken from a discussion between the boy and his father. They saw a poster for a film, where a man is standing on a whale and trying to harpoon it. The whale's size is exaggerated for the sake of sensation. The father asked what was wrong with the picture and the boy eventually said: "I know what you mean. That whale should be smaller. When we were in England we saw an orca and it was only as tall as three men" (van den Brink & Streefland, 1979, p. 405). Quoting Bryant (1974), van den Brink and Streefland argued that visual proportions are part of the basic mechanisms of perception, which can be used in learning in a variety of situations, and suggested that this might be an excellent start for making children aware of relations between quantities.

Van den Brink and Streefland then developed classroom activities where 6- to 8-year-old children explored proportional relations in drawings. The teacher showed the children a picture of a house and asked them to mark their own height on the door of the house. The children engaged in measurements of themselves and the door of the classroom in order to transpose this size relation to the drawing and mark their heights on the door. This activity generated discussions relevant to the question of proportions, but these were treated only intuitively. The lesson ended with a surprise: the teacher unveiled another part of the same picture, which showed a girl standing next to the house. The girl was much taller than the house and the children concluded that the house in the picture must be a doll house. Surprise and playfulness were considered by Streefland to be an important factor in children's engagement in mathematics lessons. Van den Brink and Streefland suggested that children can use perceptual mechanisms to reflect about proportions when they judge something to be out of proportion in a picture. They argued that it is not only of psychological interest but also of mathematical–didactical interest to discover why children can reason in ratio and proportion terms in such situations, abstracting from perceptual mechanisms.

Streefland (1984) later developed further activities in a lesson series with the theme "with a giant's regard," which started with activities that explored the children's informal sense of proportions and progressively included mathematical representations in the lesson. The children were asked, for example, to imagine how many steps would a normal man take to catch up with one of the giant's steps; later, they were asked to represent the man's and the giant's steps on a line and subsequently by means of a table. Figure 3.5 presents an example of the type of diagram used for a visual comparison.

In a later paper, Streefland (1985) pursued this theme further and illustrated how the diagrams originally used to represent visual meanings could later be used in a progressively more abstract way to represent correspondences between values in problems that did not have a visual basis. This was illustrated using, among others, Hart's (1981) onion soup problem described earlier. The diagram proposed by Streefland, which the teacher should encourage the pupils to construct, shows both (a) the correspondences between the values, which the children can find using their own informal building-up strategies, and (b) the value of the transformations that they make at every step. See Figure 3.6 for an example.

Streefland suggested that these schematic representations could be used later in vertical orientation, more common for tables than the above diagram, and with all ingredients listed on the same table in different rows. This would help students realize that when the same transformation is applied to all the ingredients, the ratio between the ingredients remains the same: in Figure 3.6, the number of people the recipe is for

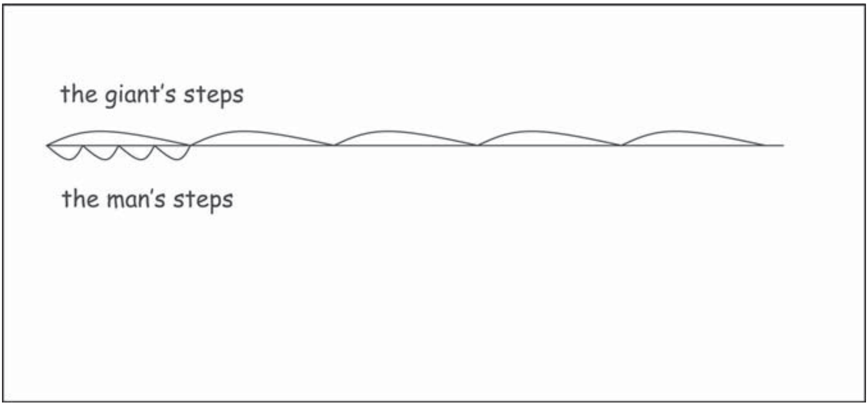


Figure 3.5 A diagram that can be used by children to represent the correspondences between a giant's and a man's steps.

Adapted from "Search for the Roots of Ratio: Some Thoughts on the Long Term Learning Process (Towards . . . a Theory): Part I: Reflections on a Teaching Experiment," by L. Streefland, 1984. *Educational Studies in Mathematics, 15,* 327–348. Reprinted with permission.

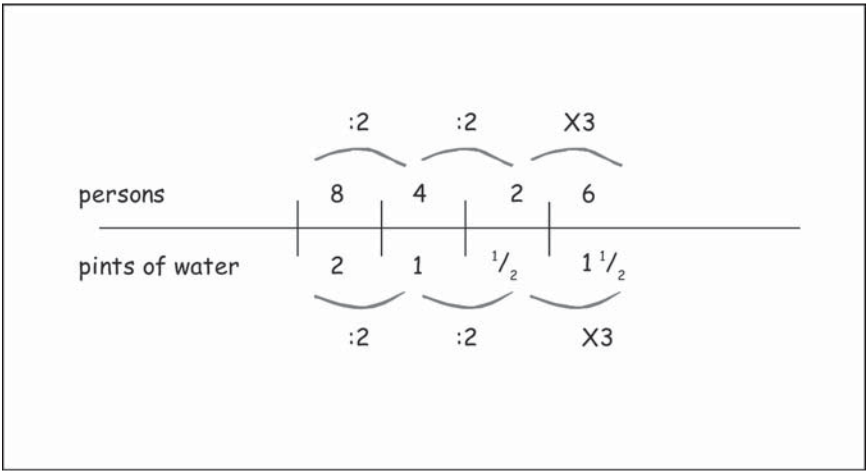


Figure 3.6 The ratio table illustrating how the building-up strategy can be represented and explored in the classroom.

Adapted from "Search for the Roots of Ratio: Some Thoughts on the Long Term Learning Process (Towards . . . a Theory): Part II: The Outline of the Long Term Learning Process," by L. Streefland, 1985. *Educational Studies in Mathematics, 16,* 75–94. Reprinted with permission.

is 4 times the amount of water across the different numbers of people being considered. Streefland (1985) argued that

> the ratio table is a permanent record of proportion as an equivalence relation, and in this way contributes to acquiring the correct concept. Applying the ratio table contributes to the detachment from the context . . . In this quality the ratio table is, as it were, a unifying model for a variety of ratio contexts, as well as for the various manifestations of ratio . . . The ratio table can contribute to discovering, making conscious and applying all properties that characterize ratio-preserving mappings and to their use in numerical problems. (p. 91)

Ratio tables should then be related to graphs, where the relation between two variables can be discussed in a new way (see also Van den Heuvel-Panhuizen, 2003).

It is noteworthy that the ideas about how intuitions about proportions can be formalized in the classroom have been explored for more than two decades now. However, there is no systematic assessment of the effectiveness of this approach in the classroom. The relative success of students in the Netherlands in international comparisons of mathematics performance (according to Bramald, 2000, they are at the top of the European countries) has not prompted experts in other countries to try the approach, although it is now widely used in both the United Kingdom and the United States.

It is most urgent that research to assess the effectiveness of this way of teaching proportions should be carried out if we want our students to go beyond the currently modest rates of success that they show in solving proportions problems. And it is also significant that the Dutch empty-number line, used to teach calculations, is so widely used but not the resources that were designed to help students go beyond arithmetic and start to understand mathematical models.

SUMMARY AND CONCLUSIONS

We started this chapter with the statement that in everyday life people find themselves having to deal with many situations in which they need to reason about proportional relations. Yet students in school and in international comparison studies show only moderate rates of success in solving proportions problems. Despite having been taught general methods for solving proportions problems, many students do not seem to use what they were taught. Instead, they resort to untaught methods, known as intuitive or informal, which are sometimes awkward and have to be implemented through a series of steps, when the use of a formal method would have led to a much simpler calculation procedure.

We reviewed research showing that unschooled adults, who have to develop their understanding of proportional relations outside of school, solve problems using much the same methods. Building-up strategies, as these informal methods have been termed, seem to have their origin in the schema of one-to-many correspondence. Some young children, aged 5 and 6 years, already show the ability to use one-to-many correspondences to solve multiplicative reasoning problems. This finding suggests that the origin of informal methods is indeed the schema of one-to-many correspondence. Finally, research has also shown that students who start school with a greater capacity to use one-to-many correspondence to solve problems achieve more in mathematics according to school standards than those with a weaker capacity to do so. Researchers in both the United States and the Netherlands have suggested ways in which children's informal knowledge could be capitalized in school in order to promote their better understanding of proportions.

It remains for us to reflect on why such research has had less impact on teaching developments than other work by the same group of researchers on arithmetic. We speculate that perhaps the focus on arithmetic goes hand in hand with the taking for granted of children's understanding of relations. The development of an understanding of relations between quantities could be seen very much as something that should happen in everyday life, not in the mathematics classroom. But, as Thompson (1994) pointed out, understanding relations is at the heart of quantitative reasoning. Freudenthal (1983) and Vergnaud (1979) argued about three decades ago that we need to know about children's implicit mathematical models for problem situations, not just their arithmetic skills, when we want to develop their problem-solving ability. "Different properties, almost equivalent to the mathematician, are not all equivalent for the child" (Vergnaud, 1979, p. 264), and problems that are solved by the same arithmetic operations have very different levels of difficulty depending on the relations that children must understand in order to solve the problems.

We therefore conclude that if mathematics education policies ignore the importance of relations and children's implicit models for situations, they do so at the risk of holding back many children's mathematics development.

ACKNOWLEDGMENTS

We are thankful to the ESRC-Teaching and Learning Research Programme, whose generous support (through grant number L139251015) made some of the work reported here possible. We are also very grateful to The Nuffield Foundation for the support it offered us to conduct a review of the literature on mathematics learning, partially reported in this chapter. The opinions expressed in this chapter are our own and not to be attributed to these institutions.

REFERENCES

Adjiage, R., & Pluvinage, F. (2007). An experiment in teaching ratio and proportion. *Educational Studies in Mathematics, 65,* 149–175.

Ainley, J., Pratt, D., & Nardi, E. (2001). Normalising: Children's activity to construct meanings for trend. *Educational Studies in Mathematics, 45,* 131–146.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement, TIMSS International Study Center, Boston College.

Becker, J. (1993). Young children's numerical use of number words: Counting in many-to-one situations. *Developmental Psychology, 19,* 458–465.

Booth, L. R. (1981). Child-methods in secondary mathematics. *Educational Studies in Mathematics, 12,* 29–41.

Bramald, R. (2000). Introducing the empty number line. *Education 3–13, 28,* 5–12.

Bryant, P. (1974). *Perception and understanding in young children.* London: Methuen.

Bryant, P. E., & Bradley, L. (1985) *Children's reading problems.* Oxford: Blackwell.

Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., & Weisbeck, L. (1993). Models of problem solving: A study of kindergarten children's problem-solving processes. *Journal for Research in Mathematics Education, 24,* 428–441.

Confrey, J. (1994). Splitting, similarity and rate of change: A new approach to multiplication and exponential functions. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 293–332). Albany: State University of New York Press.

Cramer, K., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. T. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 159–178). New York: Macmillan.

Elliott, C. D., Smith, P., & McCulloch, K. (1997). *British Ability Scale-II technical manual*. London: Nelson Publishing Company.

Fischbein, E. (1987). *Intuition in science and mathematics: An educational approach.* Dordrecht, The Netherlands: Reidel.

Flavell, J. H. (1971). What is memory development the development of? *Human Development, 14,* 272–278.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1990). Developmental changes in young children's knowledge about the mind. *Cognitive Development, 5,* 1–27.

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3,* 413–435.

Freudenthal, H. (1983). *Didactical Pheromaology of Mathematical Structures.* Dordrecht: D. Reidel.

Frydman, O., & Bryant, P. E. (1988). Sharing and the understanding of number equivalence by young children. *Cognitive Development, 3,* 323–339.

Gombert, J. E. (1992). *Metalinguistic development.* Hemel Hempstead, UK: Harverster Wheatsheaf.

Hart, K. (1981). Ratio and proportion. In K. Hart (Ed.), *Children's understanding of mathematics: 11–16* (pp. 88–101). London: Murray.

Hart, K. (1984). *Ratio: Children's strategies and errors. A report of the strategies and errors in secondary mathematics project.* Windsor, UK: NFER-Nelson.

Inhelder, B., &. Piaget, J. (1959). *The growth of logical thinking from childhood to adolescence.* London: Routledge.

Kamii, M. (1981). Children's ideas about written number. *Topics in Learning and Learning Disabilities, 1,* 47–59.

Kaput, J., & West, M. M. (1994). Missing-value proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 237–292). Albany: State University of New York Press.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

Karmiloff-Smith, A., & Inhelder, B. (1977). If you want to get ahead, get a theory. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 293–306). Cambridge, UK: Cambridge University Press.

Karplus, R., & Peterson, R. W. (1970). Intellectual development beyond elementary school II: Ratio, a survey. *School-Science and Mathematics, 70,* 813–820.

Karplus, R., Pulos, S., & Stage, E. K. (1983). Proportional reasoning of early adolescents. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 45–90). London: Academic Press.

Kieren, T. (1994). Multiple views of multiplicative structures. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 389–400). Albany: State University of New York Press.

Kouba, V. (1989). Children's solution strategies for equivalent set multiplication and division word problems. *Journal for Research in Mathematics Education, 20,* 147–158.

Kramarski, B. (2004). Making sense of graphs: Does metacognitive instruction make a difference on students' mathematical conceptions and alternative conceptions? *Learning and Instruction, 14,* 593–619.

Lamon, S. (1994). *Ratio and proportion: Cognitive foundations in unitizing and norming.* In G. Harel & J. Confrey (Eds.), The development of multiplicative reasoning

in the learning of mathematics (pp. 89–122). Albany: State University of New York Press.
Lesh, R., Post, T., & Behr, M. (1988). Proportional reasoning. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 93–118). Reston, VA: National Council of Teachers of Mathematics.
Noelting, G. (1980a). The development of proportional reasoning and the ratio concept Part I—Differentiation of stages. *Educational Studies in Mathematics, 11,* 217–253.
Noelting, G. (1980b). The development of proportional reasoning and the ratio concept Part II—Problem-structure at successive stages: Problem-solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics, 11,* 331–363.
Nunes Carraher, T. (1986). From drawings to buildings: Working with mathematical scales. *International Journal of Behavioral Development, 9,* 527–544.
Nunes, T., & Bryant, P. (1996). *Children doing mathematics.* Oxford: Blackwell.
Nunes, T., & Bryant, P. (2009). *Children's reading and spelling. Beyond the first steps.* London: Wiley/Blackwell.
Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. N. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology, 25,* 147–166.
Nunes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics.* New York: Cambridge University Press.
Piaget, J. (1952). *The child's conception of number.* London: Routledge.
Piaget, J. (1978). *The grasp of consciousness.* London: Routledge.
Piaget, J. (2001). *Studies in reflecting abstraction: Translated by R. Campbell.* Hove, UK: Psychology Press.
Piaget, J. (2008). Intellectual evolution from adolescence to adulthood. *Human Development, 51*, 40–47 (Reprint of *Human Development* 1972;15:1–12).
Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* London: Routledge.
Pickering, S., & Gathercole, S. (2001). *Working memory test battery for children (WMTB-C) manual*. London: The Psychological Corporation.
Resnick, L. B. (1983). Mathematics and science learning: A new conception. *Science, 220,* 477–478.
Ricco, G. (1982). Les première acquisitions de la notion de fonction linèaire chez l'enfant de 7 à 11 ans. *Educational Studies in Mathematics, 13,* 289–327.
Rojas-Drummod, S., Pérez, V., Vélez, M., Gómez, M. & Mendoza, A. (2003). Talking for reasoning among Mexican primary school children. *Learning and Instruction, 13,* 653–670.
Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition, 92,* 329–352.
Sellke, D., Behr, M., & Voelker, A. (1991). Using data tables to represent and solve multiplicative story problems. *Journal for Research in Mathematics Education, 22,* 30–38.
Steffe, L. (1994). Children's multiplying schemes. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 3–40). Albany: State University of New York Press.

Streefland, L. (1984). Search for the roots of ratio: Some thoughts on the long term learning process (towards . . . a theory): Part I: Reflections on a teaching experiment. *Educational Studies in Mathematics, 15,* 327–348.

Streefland, L. (1985). Search for the roots of ratio: Some thoughts on the long term learning process (towards . . . a theory): Part II: The outline of the long term learning process. *Educational Studies in Mathematics, 16,* 75–94.

Thompson, P. W. (1993). Quantitative reasoning, complexity, and additive structures. *Educational Studies in Mathematics, 25,* 165–208.

Thompson, P. (1994). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 181–236). Albany: State University of New York Press.

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. *Educational Studies in Mathematics, 16,* 181–204.

Treffers, A. (1987). Integrated column arithmetic according to progressive schematisation. *Educational Studies in Mathematics, 18,* 125–145.

Van den Brink, J., & Streefland, L. (1979). Young children (6–8): Ratio and proportion. *Educational Studies in Mathematics, 10,* 403–420.

Van de Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. Realistic mathematics education research: Leen Streefland's work continues. *Educational Studies in Mathematics, 54,* 9–35.

Vergnaud, G. (1979). The acquisition of arithmetical concepts. *Educational Studies in Mathematics, 10,* 263–274.

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 128–175). London: Academic Press.

Vergnaud, G. (1997). The nature of mathematical concepts. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 1–28). Hove, UK: Psychology Press.

Vergnaud, G. (in press). The theory of conceptual fields. *Human Development*.

This page intentionally left blank

4 Students' Goals Influence Their Learning

XIAODONG LIN, ROBERT S. SIEGLER,
AND FLORENCE R. SULLIVAN

Research on learning occupies a prominent place in both psychology and education. Psychological research on learning usually focuses on specific learning processes: analogical reasoning, causal inference, discrimination, encoding, retrieval, formation of mental representations, strategy formation, and problem solving (Bransford, Brown, & Cocking, 1999; Brown, Bransford, Ferrara, & Campione, 1983; Siegler, 2006). Methods such as expert–novice comparisons and microgenetic designs have yielded valuable lessons regarding the differences produced by deep mastery of specific content and the variability of learning processes both within and between learners (Chase & Simon, 1973; Kuhn & Franklin, 2006; Siegler, 2006; Staszewski, 1988).

Research on learning has also been conducted by researchers primarily interested in education. Educationally oriented studies of learning often emphasize how to promote deeper learning and reflection through instructional scaffolds, such as designing instructional methods and technologies and comparing the quantity and quality of learning that they elicit (Bransford et al., 1999). The rapid growth of technology in recent years, particularly the connectivity of the Internet and television, has prompted educational researchers to focus on dynamic relations among learners, instructors, instructional content, and instructional media (Cognition & Technology Group at Vanderbilt, 1997; Lin, 2001; Lin & Lehman, 1999).

A general assumption underlying most studies of learning in both psychology and education is that the primary goal of learners is to master the material. Although this assumption might seem obvious, our research indicates that it probably is not valid. Learners' primary goals, at least in the school context, are highly variable, and deep understanding of the material is not always a high priority (Lin, Schwartz, & Hatano, 2005). Instead, learners' goals vary with their social and cultural contexts, including the school context. Instructional techniques and environments that recognize this diversity of goals and that address it can increase learning.

In this chapter, we first describe several studies showing that not all students view learning as their primary goal in school and that students' goals vary between and within cultures. Next, we discuss the effects on students' understanding of how an instructor's background influenced her goals and values, on students' attitudes toward the instructor, and on the quality of collaborative teacher–student problem solving. Finally, we explore how educational approaches that encourage students to adopt the goal of explaining the behavior of other people and of physical devices can improve learning. The overarching message is that the optimization of students' learning requires attention to learner's goals and values, and that instruction which leads students to adopt the goal of deeply understanding the material that is being taught can produce superior learning.

IS LEARNING THE PRIMARY GOAL FOR ALL STUDENTS?

To examine whether learning is students' primary goal in school, Lin and Schwartz (2007) asked 371 fifth-grade students and their 12 teachers to design an ideal student. That is, each fifth grader and teacher was asked to select the five properties that they thought were most important for an ideal student. The students and teachers came from 12 classes in three schools in New York City (four classes per school). School A was in a neighborhood of Harlem that served a predominantly low-income African American population. School B was in a neighborhood in Chinatown that served a predominately low-income East Asian population. School C was in a neighborhood in midtown that served a population of mixed socioeconomic and ethnic backgrounds. As shown in Figure 4.1, school A had the lowest average achievement test scores on both the ELA (a measure of reading skills in English) and a math achievement test, school B had the highest average achievement on both tests, and school C was between on both.

We coded and analyzed the ideal-student properties generated by each student and teacher in this free-response format. The properties that were generated fell into four major categories: learning well, behaving well, performing well, and socializing well. In the present discussion, we mainly focus on the learning and behavior categories, because these are the areas where the mismatches between teachers' and students' goals were most pronounced. Responses that emphasized learning included "being able to understand what is taught," "knowing when he/she makes mistakes," "asking good questions," "explaining ideas clearly," and so on. Responses that emphasized behavior included "not fighting in class," "sitting still when the teacher is lecturing," "raising hands before answering questions," "following class rules," and so on.

Almost all children at all schools generated at least one aspect of good behavior as an important characteristic of an ideal student. However, as shown in Figure 4.2, the percentage of students from the three schools who chose at least one aspect of good learning as an important goal for their ideal student was far more variable. The school with the highest academic achievement (school B) had almost twice as high a percentage of students who considered high-quality learning as an important characteristic of an ideal student as compared to the school with

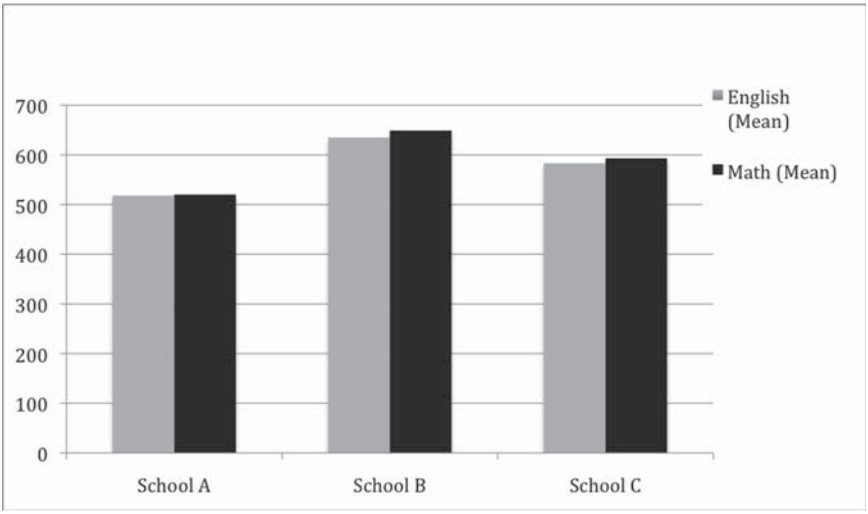


Figure 4.1 Mean English and mathematics achievement test scores for three schools.

the lowest academic achievement (school A)—61% versus 32%. The school with an intermediate level of academic achievement also had an intermediate level of students citing at least one learning goal for their ideal student (45%).

Analyses at the level of individual classrooms yielded converging evidence for this analysis. As shown in Figure 4.3, the highest-achieving class had a much higher percentage of students citing learning goals than the lowest achieving class: 62% versus 26%. Regardless of ethnicity, school types, and socioeconomic status (SES), students with higher academic achievement tended to assign more learning-related properties to their ideal student than did low-achieving students.

Another question of considerable interest was whether students and teachers have different goals for their ideal student. Teachers at all of the schools were far more likely than their students to emphasize learning goals. Almost all teachers—75% of those from school A and 100% of those from schools B and C–included learning goals among the properties of their ideal student. Thus the teachers' goals were not well aligned with the students' goals in that the teachers put far more emphasis on learning goals, even at the school with the lowest achievement.

At the level of individual classrooms, 9 of the 12 classes had a significant mismatch between how the teacher and his or her students

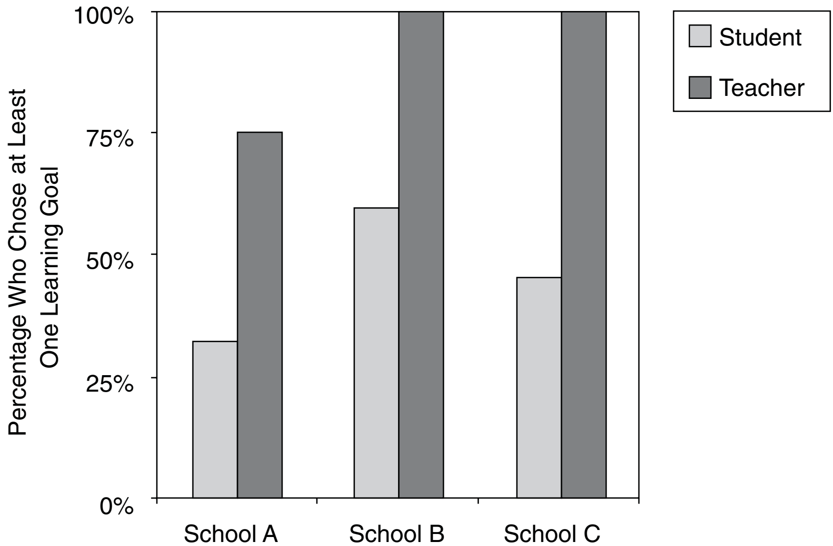


Figure 4.2 Percentages of students and teachers from the three schools who chose at least one aspect of good learning as an important goal for their ideal student.

envisioned an ideal student. The teachers generated almost twice as many learning-oriented qualities for their ideal student as the students. In contrast, the students generated far more qualities relevant to good classroom behaviors and attractive personalities.

We compared the degree of mismatch between students' and teachers' ideal-student characteristics of the two classes that had the highest achievement scores with those of the two classes that had the lowest achievement scores. The mismatch was larger for the low-achieving classes than for the higher-achieving classes (Figure 4.4). Teachers in all four classrooms cited one or more learning goals. The same was true for almost all students in the highest-achieving classrooms (almost 90%). In contrast, only about half of the students in the lowest-achieving classrooms cited any learning goals.

Mismatches between teachers' and students' goals may be one reason why innovative materials and instruction have not fostered the learning of all students (Cuevas, Lee, Hart, & Deaktor, 2005; Lee, 2003; 2004; Lin & Schwartz, 2003). Based on our results to date, many teachers believe that students' major goal is to learn academic content and acquire skills. However, the present data indicate that many students in urban communities do not share this goal or at least do not view it as a high priority. Instructional programs may be more effective if students assign learning goals a higher priority, so that their goals, those of

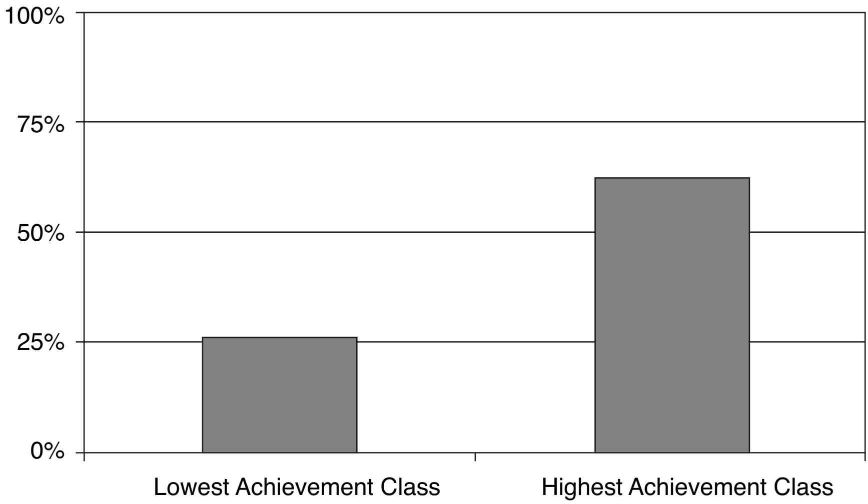


Figure 4.3 Comparison of percentages of students in the highest and the lowest achieving classes who generated learning goals for their ideal student.

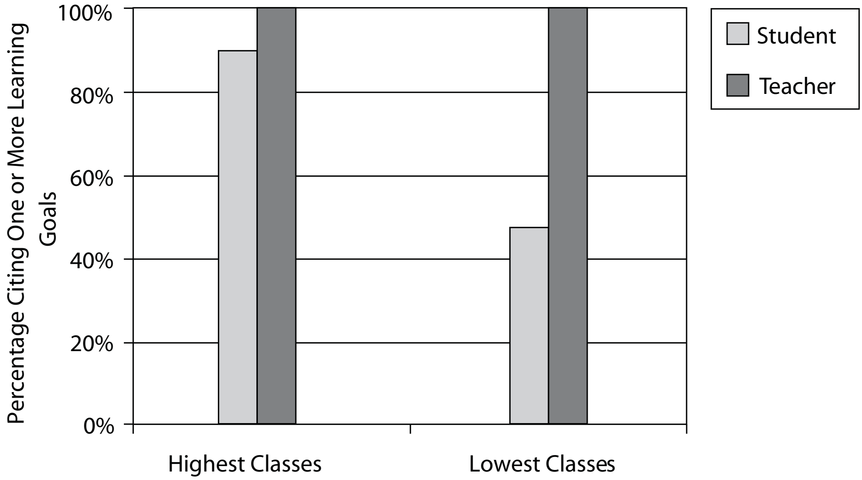


Figure 4.4 Comparison of students and teachers in the highest and lowest discipline and achieving classes who cited one or more learning goals.

teachers, and those implicitly valued in the entire academic system are more closely aligned. Perhaps because educators have assumed that everyone knows that the goal of school is to build knowledge and that everyone adheres to this goal, schools have failed to inculcate this goal in all of their students.

CROSS-NATIONAL DIFFERENCES IN LEARNING GOALS

The findings summarized above indicate that students' goals and values vary greatly among classes and individual students in the United States. We have also examined whether students' and teachers' ideal students vary between cultures. In this study, we asked 280 fifth-grade students and their teachers in both public and private schools in China and in the United States to specify the five most important characteristics of an ideal student (Lin & Schwartz, 2003). The methodology was very similar to that used in the study that contrasted the three schools in New York City.

Figure 4.5 presents the percentage of students and teachers in the United States and China who generated one or more learning characteristics and one or more aspect of good behavior for their ideal student. As in the study of New York City schools, virtually all teachers cited learning goals. The Chinese students in both public and private schools and

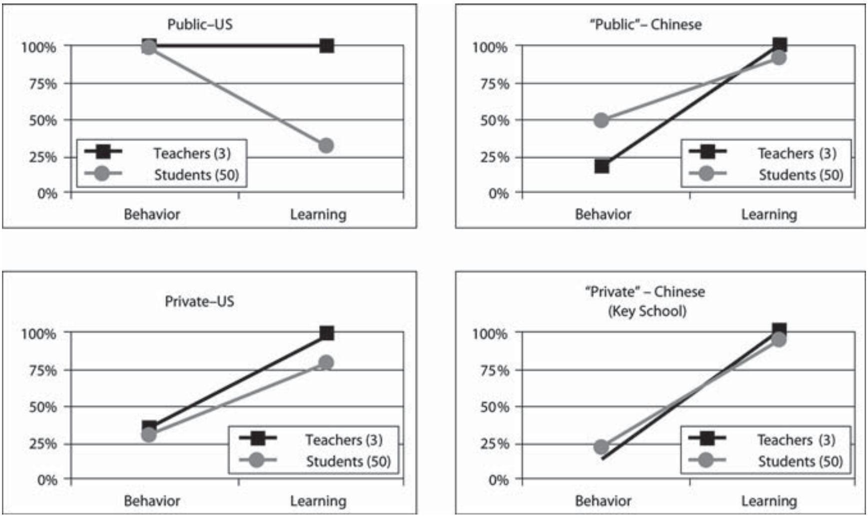


Figure 4.5 Comparisons of ideal student characteristics generated by U.S. and Chinese public and private schools, teachers, and students.

the American students in private schools also emphasized learning goals (e.g., "the ideal student explains and understands deeply and tries hard to correct mistakes when they occur"). In contrast, the American public school students were concerned mostly with good classroom behavior (e.g., "does not fight, sits still during lectures"). Apparently, students' goals differ both within cultures (U.S. public and private schools) and between cultures (U.S. and Chinese public schools). It seems clear from Figure 4.5 that some parents, teachers, and schools in the United States as well as in China had communicated successfully to their students that learning well was essential to being a good student.

Significant for this discussion, the American public school teachers, like their students, emphasized good behavior in their choices of properties of ideal students far more than the other groups of teachers. Good behavior is far more of an issue in the American public schools than in the Chinese public schools or in the private schools in both countries. The effect on both students and teachers is to lead them to cite good behavior more often as a characteristic of the ideal student. However, the American public school teachers also cite learning goals as essential, while this goal seems to be lost on many students.

We showed these results to teachers in the U.S. public schools who had filled out the questionnaires regarding the ideal student. The teachers initially expressed surprise at their students' emphasis on good behavior at the expense of learning well. Then they noticed that they, too, were emphasizing behavior more than did the teachers from other schools. The teachers then began to reflect on their own assumptions about teaching, their actions in the classroom, how they could improve their practices, and how they could communicate the centrality of learning to their students.

These findings suggest that the common assumption that all students view learning as their primary goal in attending school is false. Students' goals and values vary greatly depending on the families, teachers, schools, and local cultures within which they live and interact. The results suggest that learning research should include examination of students' goals and values regarding schooling, how teachers can be made aware of their own and their students' values, and how teachers can modify their behaviors so that their students adopt the learning of academic material as an essential goal for themselves.

DOES UNDERSTANDING PEOPLE'S GOALS AND VALUES IMPROVE CROSS-CULTURAL UNDERSTANDING?

If we take seriously the proposition that culture greatly influences goals and values, then the culture of the classroom should influence students' goals and values for learning (Cole, 2002; Nisbett, Peng, Choi, & Norenzayan, 2001; Rogoff, 2003). For example, Rosenthal and Jacobson (1968) concluded that students' performance reflects teachers' expectations; when teachers set more ambitious goals for their students, students learn more than when teachers present them with less ambitious goals. Yet teachers are often unaware of their own values and beliefs (Cohen, 1991; Darling-Hammond, 1997; Lampert, 1990; Staub & Stern, 2002; Thompson, 1992) and find it difficult to describe them (Jacobs, Yoshida, Stigler, & Fernandez, 1997; Jacobs & Morita, 2002). As the saying goes, "Fish will be the last to discover water." This may be one of the reasons why teachers find it difficult to revamp their traditional methods of instruction (Darling-Hammond, 1997; Prawat, 1992; Thompson, 1992). When teachers do not have an explicit understanding of their own goals, it is difficult for them to see how those goals can best be implemented, much less how to change these goals and practices.

In this section, we address the question of whether gaining insight into other people's goals and cultural backgrounds improves our ability to solve problems involving intercultural working relationships. To answer this question, we examined how two types of video-based stories that contained personal background knowledge (PBK) and general cultural background knowledge (GCBK) affected people's understanding of an educational problem and proposed ways of solving it. The problem-solving situation involved a poor relationship between a foreign-born college professor and her students (Lin & Bransford, in press).

The GCBK video included general information about the culture in which the professor grew up: its customs, history, and political and social systems. This video is representative of the ways in which past research has tried to bridge cultural gaps in understanding (for example, see Eini, 2006).

The PBK video represents a very different genre of communication: one that focuses on the individual teacher's goals, values, and formative experiences. Because of this emphasis, it was hypothesized to offer more powerful ways of helping students understand and empathize with the professor and thus motivate them to perform better in her class. The PBK included detailed information regarding the professor's personal experiences and upbringing, the goals and values that she and her family viewed as most important, and why she and her family viewed those goals and values as most important. It also linked this background to her approach to teaching in order to help students understand why she taught the way she did.

The participants were 43 preservice teachers (25 women and 18 men) enrolled in a general educational psychology course. The problem case used in the study was inspired by a real classroom experience that evolved into a highly uncomfortable conflict between a foreign-born college professor and her U.S. college students. As the conflict intensified, many students began to complain to fellow students and other professors that they had been stuck with a foreign-born professor with an accent who assigned too many readings, graded too strictly, and even took attendance. The professor, a new PhD, was trying her best to give the students the highest-quality education she could provide. Over the course of the semester, the sense of disconnection between the professor and her students worsened, and the professor began to search for ways to start afresh and reverse the downward spiral in her students' morale.

A review of the literature, as well as discussions with other foreign-born professors, suggested that this problem was not unique to this particular professor (Alberts, 2008; Ngwainmbi, 2006). One sign of the pervasiveness of the problem is that many university teaching centers have programs to address exactly this kind of issue. The programs that we found focused primarily on helping new professors adjust to American culture. This approach seemed likely to be helpful but also limited, because it placed little emphasis on helping university students rethink their attitudes about professors (and other people) who were different from them and who came from foreign cultures.

Lin and Bransford (in press) created an experimental version of the student–teacher "disconnect" phenomenon described above. The first step was to build a multimedia casebook that described the behaviors of "Professor X." She was said to assign too much homework and demand too much detail in papers. Included in the casebook was a wide-ranging set of negative statements from students; these mirrored many that had been expressed in the actual case that motivated this study. The casebook, which included video recordings, set the stage for exploring how different treatment conditions might affect participants' subsequent attitudes and questions about Professor X as well as their strategies for resolving the conflict.

The study combined a within-and-between-subjects design. The participants first heard, read, and voiced their opinions about Professor X and her students without having access to any information about either the personal or impersonal cultural background of the professor (baseline measure). Then, the participants were assigned randomly to one of two conditions that differed in the video the students were shown: Personal Background Knowledge (PBK) or General Cultural Background Knowledge (GCBK). The PBK video started with a narrative about how Professor X and her family were affected by the Cultural Revolution in China. Students learned that the entire family was sent from the large city in which they had previously lived a privileged existence to a poor, remote rural area where they lived in a cave, were deprived of books, and had no teachers, formal classes, or other educational opportunities. Because of this experience, Professor X particularly valued educational opportunities, and she adopted learning as her primary goal when, after the Cultural Revolution had ended, she was given the opportunity to attend a university. The video also described the great efforts that Professor X, her family, and the other people who had been exiled made to educate themselves and each other during those hard times. This experience led her to value education strongly and to react negatively when her students seemed to take college education for granted and refused to apply themselves to learning. Her view was that students should make learning their primary goal, relegating parties and other social activities to the background.

Participants in the study who were assigned to the GCBK condition were given general, impersonal information about China. They were told about Chinese history; how Mao copied the first emperor of China to start the Chinese Cultural Revolution; how the ancient Chinese developed their language, political systems, and food; and how modern Chinese celebrate various holidays. There was no information in the GCBK condition about individual Chinese, how they and their families lived through the Cultural Revolution, or how it affected their goals and values. After the students read and watched the PBK or the GCBK video, the researchers collected and analyzed the participants' assessment of Professor X's personality, their explanations of the causes of the conflict between Professor X and her students, and their proposed solutions for resolving the conflict.

The PBK video had strong positive effects on students' understanding and interpretation of the problem situation and on their strategies for solving it. In contrast, the GCBK video tended to worsen negative stereotypes and opinions of both Professor X and Chinese society. This latter outcome was surprising, given the frequent reliance on such videos and the background assumption that general cultural knowledge makes people more empathetic for those from dissimilar backgrounds. If these negative effects prove to be general, the usual general information videos that are used in efforts to build cultural understanding will need to be seriously reconsidered.

Prior to viewing the videotapes, almost all of the students saw the problem as being caused by Professor X and her unrealistic expectations. Students who watched the GCBK video did not change their perceptions. For example, one student wrote, "The professor is a typical Chinese who is rigid, critical, and boring." Another student wrote, "Like most Chinese, she is hard-working and values education but is boring and strict and has few social skills." In contrast, students who watched the PBK—which told about the experiences that influenced Professor X's personal goals and values—altered their thinking. In particular, they integrated Professor X's cultural experiences into their understanding. One student wrote, "The professor realizes what life can be like without education because of her personal cultural experiences. She is a responsible professor, values education, and wants to provide her students with a good education."

To assess the degree to which the two videos influenced the students' understanding of the problems posed by the teacher–student interactions, we asked participants from both conditions to rate the change in their understanding of the problem situation. Our assumption was that if they discerned changes, they would be in a better position to reflect on those changes. At the end of the study, students rated their level of understanding before and after the videotape on a scale of 1 to 5. Both groups of students rated their initial understanding at an average level of 2.1. However, students who were given the personal background information rated their subsequent understanding at an average of 4.2, whereas the students who were given the general cultural background information provided an average rating of 2.4, barely different from the baseline.

Viewing the PBK video also improved the quality of solutions that the students offered to deal with the conflict between Professor X and her students. Relative to both their own baseline performance and the posttest performance of students who watched the GCBK video, students in the PBK condition generated better solutions after watching the video. Their solutions were more detailed and included more specific ideas that could be adopted, such as having both the teacher and the students discuss their backgrounds to improve their understanding of why they emphasized particular educational values and goals. The solutions offered by those who had watched the PBK video were also more likely to include components that would please both the professor and the students rather than being based entirely on the perspective of one side or the other.

The results of viewing the PBK video suggest that knowledge about goals and values can offer teachers and students new explanations for phenomena that are difficult to understand without such knowledge. For instance, in the ideal-student project, the U.S. public school teachers from classrooms with low achievement test scores told us that they did not understand why their implementation of innovative science activities, such as problem-based learning, did not interest their students nearly as much as expected. These innovative science curricula also did not lead to the same increases in learning in their classrooms as in other classrooms described in the literature. These teachers thought that good learning programs should interest and engage all students and that the students who were not engaged must just be dumb, lazy, or both. After participating in the ideal-student study, they realized that neither of these explanations might be right. Instead, they understood that many of their students might not have viewed learning as being among the primary goals of going to school. This knowledge gave the teachers opportunities to revise their instruction in ways that would influence students' goals, with the aim of improving their learning.

These findings are consistent with results from the previously described ideal-student studies. In both cases, students' goals and values differed from expectations, and the goals influenced the students' learning and problem solving. Also in both cases, as a result of experiences that led teachers and students to reflect on their goals, classroom interactions were influenced in positive ways.

Knowledge about students' and teachers' goals also can help us identify new areas of research and new variables for investigation. For instance, the realization that not all students view learning as their primary goal at school led us to investigate variables that we would otherwise not have considered, such as the roles of personal background knowledge in bridging classroom cultural gaps. Currently, we are exploring how explicit instruction in negotiating and handling diverse goal orientations in a science classroom can improve teacher training and whether this has consequences for student learning. We will compare classes where no information about students' and teachers' goals is provided with classes in which the teachers and students discuss each other's goals in the context of solving specific classroom problems (e.g., failure to engage in science projects; poor quality project reports, etc.). We will test how such interventions affect classroom discourse, engagement in learning, and learning outcomes. In addition, we are exploring the benefits of having teachers and students choose attributes for computer agents that participate in learning situations and observe how their ideal agents perform in various situations. A focus on learners' goals and values led us to examine the effects of these and other manipulations that may lead students to revise their goals and learn more effectively.

ENCOURAGING STUDENTS TO ADOPT EXPLANATORY GOALS

Perhaps the prototypical learning goal is striving to explain for oneself how and why events occur. As with learning goals in general, however, such explanatory goals turn out to be assigned a relatively low priority by many U.S. students. The reasonable expectation that students would consistently try to explain unexpected statements found in textbooks and heard from teachers, and unexpected observations of physical events, turns out not to be true. Yet when they do occur, such *self-explanations* enhance learning.

Self-explanations are inferences about causal connections among objects and events. The inferences can concern how procedures cause their effects, how structural aspects of a system influence its functioning, how people's reasoning leads to their conclusions, how characters' motivations within a story lead to their behavior, and so on.

The ability to infer such causal connections is present very early in life. Infants in their first year sometimes infer connections between physical causes and their effects (Leslie, 1982; Oakes & Cohen, 1995). Infants and toddlers also remember events that reflect a coherent causal sequence better than ones in which the causality is unclear (Bauer & Mandler, 1989). Thus the ability to explain the causes of events seems to be a basic property of human beings and influences many aspects of cognition, including memory, problem solving, and conceptual understanding.

Although very young children can generate causal connections, older children and adults often fail to do so. This poses a particular problem in math and science learning. Math and science teachers frequently lament the fact that their students can execute procedures but have no idea why the procedures work. Consistent with the teachers' view that such conceptual understanding is crucial, better and worse learners differ in the frequency with which they try to explain what they are learning. In a wide range of areas—including physics, biology, algebra, and computer programming—the frequency with which learners try to explain the logic underlying statements in a textbook is positively related to their ability to learn the material covered in the textbook (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, deLeeuw, Chiu, & LaVancher, 1994; Ferguson-Hessler & de Jong, 1990; Nathan, Mertz, & Ryan, 1994; Pirolli & Recker, 1994).

Explaining Observations

In recent years, a number of researchers have attempted to supplement these correlational studies with experimental evidence. That is, they have examined whether encouraging randomly selected students to adopt the learning goal of explaining their observations would increase their learning. These studies have consistently shown that being asked to explain why events occur promotes the more rapid and more frequent discovery of superior rules and strategies than does making the same observations but not explaining them (Calin-Jageman & Ratner, 2005; Chi et al., 1994; Pine & Messer, 2000; Renkl, 2002; Renkl, Atkinson, Maier, & Staley, 2002; Siegler & Chen, 1998).

Requests to explain observations have positive effects even on the learning of children who are just starting school. For example, Siegler

(1995) examined whether requests to explain why other people reached the conclusions they did would increase 5-year-olds' learning of number conservation. Children were shown two parallel rows, each with the same number of objects (7, 8, or 9) arranged in 1:1 correspondence. At the beginning of each trial, children readily agreed that the two rows had the same number of objects. Then, one of the rows was lengthened, shortened, or left spatially unchanged and had an object added, subtracted, or neither. The experimenter called attention to both spatial and numerical transformations, by saying, for example, "Now I'm spreading this row out and I'm taking an object away from it." Children in all groups were then asked whether they thought the transformed row had more objects, fewer objects, or the same number of objects as the untransformed row.

At the outset of the experiment, children in all groups were given a pretest. Those whose performance indicated that they did not yet know how to solve number conservation problems were randomly assigned to one of three experimental conditions. One group of children received feedback alone; they advanced their answer and were immediately told whether it was correct or incorrect (*feedback-only condition*). A second group of children advanced their answers and were asked, "Why do you think that?" Then they were given feedback on their answers (*explain-own-reasoning condition*). Examining this condition allowed us to determine whether describing one's own reasoning was causally related to learning.

A third group of children advanced their answers, received feedback from the experimenter concerning which answer was correct, and were then asked by the experimenter "How do you think I knew that?" (*explain-correct-reasoning condition*). This last condition, in which the child was asked to explain the experimenter's reasoning, was of greatest interest. Having children explain another person's correct reasoning combines advantages of discovery and didactic approaches to instruction. It is like discovery-oriented approaches in that it requires the child to generate a relatively deep analysis of a phenomenon without being told how to do so. It is like didactic approaches in that it focuses the child's attention on correct reasoning. Thus it combines some of the efficiency of didactic instruction with some of the motivating properties of discovery.

The results indicated that, as hypothesized, encouraging children to explain the reasoning underlying the experimenter's answer resulted in their learning more than feedback alone or feedback in combination

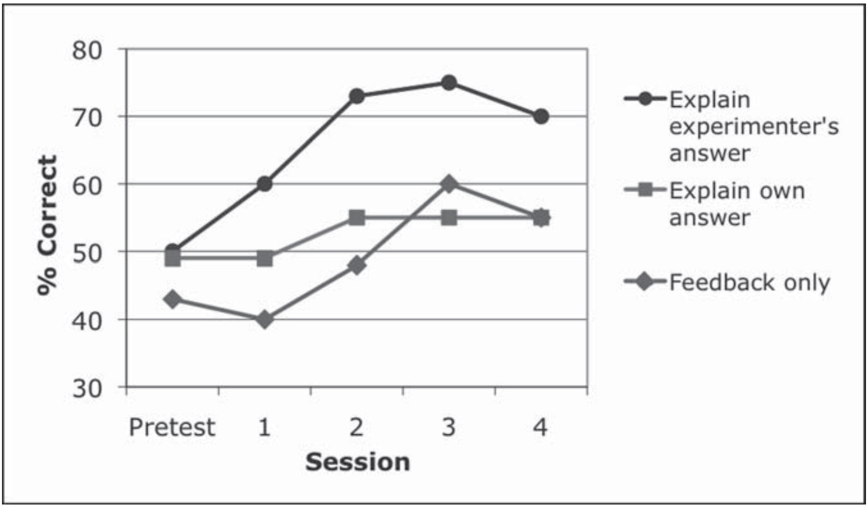


Figure 4.6 Percentages of students with correct reasoning on a number conservation task during pretest and training sessions.

with requests to explain their own reasoning (Figure 4.6). The differential gains were largest on the most difficult problems—those in which relying on the length cue led to the wrong answer. Those children who explained the experimenter's judgment in terms of the numerical transformation that had been performed learned far more than children who generated other types of explanations or those who could not generate any explanation for the experimenter's judgment.

Explaining Both Correct and Incorrect Answers

Other studies have examined the effects of encouraging learners to explain why wrong answers are wrong as well as why right answers are right. Within recent computer simulation models of strategy choice, such as ASCM and SCADS (Shrager & Siegler, 1998; Siegler & Shipley, 1995), the likelihood of a strategy being used on a problem is a positive function of its own effectiveness and a negative function of the effectiveness of competing approaches. For example, although children can solve 2 + 2 very quickly and accurately by counting from 1, they rarely use that approach because they can solve 2 + 2 even more quickly and just as accurately by retrieving the answer from memory. Similarly, a strategy that is not particularly fast and accurate will be used often if alternative approaches are even less effective. Thus the likelihood of using a given strategy can be increased in two ways: increasing its own strength or decreasing the strength of alternative strategies.

The computer simulations suggest that the best way to increase the use of the new, more advanced approaches should be to increase their strength and also to decrease the strength of less advanced approaches. In the context of self-explanation, having children explain both why correct approaches are correct and why incorrect approaches are incorrect should be more effective than explaining only why correct approaches are correct. Explaining how correct answers were generated and why they are correct should increase the strength of correct procedures; explaining how incorrect answers were generated and why they are wrong should decrease the strength of incorrect procedures.

Siegler (2002) tested this prediction on the mathematical equality task developed by Perry, Church, and Goldin-Meadow (1988). This task involves problems of the form A + B + C = ____ + C. Third- and fourth-graders find such problems surprisingly difficult. For example, they usually answer 3 + 4 + 5 = ____ + 5 by writing "12." This answer reflects an add-to-equal-sign strategy, in which the children add all numbers to the left of the equal sign. The next most common answer to the problem is 17, which reflects an add-all-numbers strategy. Both approaches reflect limited understanding of what the equal sign means. The third and fourth graders seem to interpret it either as meaningless or as a signal to add the relevant numbers rather than as an indication that the values on the two sides of the equal sign must be made equivalent.

In Siegler (2002), 87 third- and fourth-graders were presented a procedure that included three phases: pretest, training, and posttest. The pretest and posttest included three types of problems: A + B + C = ____ + C (*C problems*), A + B + C = ____ + B (*B problems*), and A + B + C = ____ + D (*D problems*). These problems differed in the relation of the number after the equal sign to the numbers before it. On C problems, the number after the equal sign was identical to the rightmost number before it (e.g., 3 + 4 + 5 = ___ + 5). On B problems, the number after the equal sign was identical to the middle number before it (e.g., 3 + 4 + 5 = ____ + 4). On D problems, the number after the equal sign did not match any of the numbers before it (e.g., 3 + 4 + 5 = ____ + 6).

The reason for including these three kinds of problems was that they were solvable by different types of strategies that children might induce from the feedback. The strategy of just adding the first two numbers worked on C problems but not on B or D problems. The strategy of locating a number present on both sides of the equal sign and adding the other two numbers works on B and C problems but not on D problems. Two other strategies worked on all types of problems as well as implying conceptual understanding of the equal sign. One of these optimal strategies was to create equivalent values on the two sides of the equal sign (e.g., on 3 + 4 + 5 = ___ + 5, add the numbers on the left and solve 12 = ___ + 5). The other optimal strategy was to subtract from both sides the number on the right side of the equation (e.g., on 3 + 4 + 5 = ___ + 5, subtract 5 from both sides and solve 3 + 4 = ___). These two strategies would be effective on any mathematical equality problem. Thus, presenting these three types of problems made it possible to assess children's use of strategy before and after training.

The training procedure included 10 problems. Those of greatest interest were the 6 C problems, such as 3 + 4 + 5 = ____ + 5. The other 4 items were standard 3-term addition problems with no numbers on the right side of the equal sign, such as 5 + 6 + 7 = ____. These four problems were included to prevent children from developing the approach of blindly adding the first two numbers on all problems. Performance on these foils was virtually perfect in all conditions and is not described further.

Children received the 10 problems under one of three training conditions. Children in the *explain-own-reasoning condition* were asked to answer a problem and then asked to explain why they thought their answer was correct; they were then given feedback (either "You're right, the answer is N" or "Actually, the correct answer is N"). Children in the *explain-correct-reasoning condition* also were presented a problem, asked to answer it, and given feedback as to the correct answer. However, they were then told that a child at another school had answered N (the right answer), asked how they thought the child at the other school had done so, and asked why they thought that was the right answer. Finally, children in the *explain-correct-and-incorrect-reasoning condition* were presented the same procedure except that they were asked to explain not only the reasoning of a hypothetical child who had generated the right answer but also the reasoning of a hypothetical child who had generated a wrong answer. The wrong answer that children in this condition needed to explain matched the answer that would have been generated by the strategy that that child had used most often on the pretest.

As shown in Figure 4.7, children in all conditions learned a considerable amount during training. However, those who were asked to explain both why correct reasoning was correct and why incorrect reasoning was

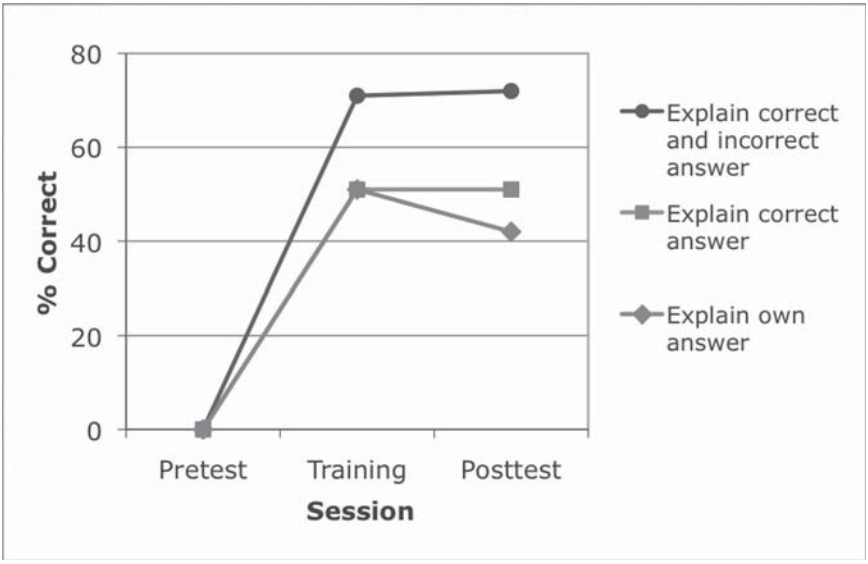


Figure 4.7 Percentages of students with correct reasoning on a mathematical equality task during pretest, training, and posttest sessions.

incorrect learned more than those in the other two groups. These differences were maintained on the posttest.

As shown in Figure 4.8, the superior posttest performance of children who explained both correct and incorrect answers during training was due largely to their being better able to solve the problems that required relatively deep understanding (B and D problems). Analysis of changes in explanations during the training phase made clear the source of this effect. Children in all groups greatly decreased their use of the add-to-equal-sign strategy, which had predominated on the pretest. The decrease occurred more quickly in the group in which children needed to explain why that strategy was wrong; but over the six trials, it occurred in all the groups to large extents. However, the groups differed considerably in the new strategies that children adopted. Children who received only feedback and explained their own reasoning largely adopted the simplest strategy, that of adding A + B. In contrast, children who explained both why correct answers were correct and why incorrect ones were incorrect were more likely to use the advanced strategies of equalizing the two sides or eliminating the constant on the right side of the equal sign by subtracting its value from both sides.

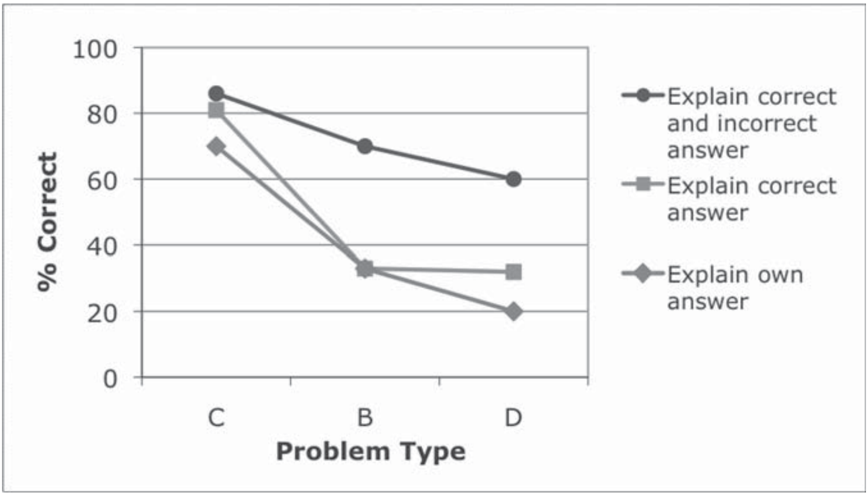


Figure 4.8 Percentages of students with posttest correct reasoning on a mathematical equality task centering on three types of problems: trained (C), near generalization (B), and far generalization (D).

The strategies that children adopted to explain correct answers during the training period proved to be very predictive of their own posttest performance. Frequency of adopting one of the two advanced strategies correlated $r = .77$ with percent correct on the B problems and $r = .86$ with percent correct on the D problems on the posttest. In contrast, percent use of the A + B explanations during training was strongly negatively correlated with performance on these problems: $r = -.70$ with percent correct on the B problems and $r = -.76$ with percent correct on the D problems on the posttest. Thus asking children to explain why correct answers were correct and why incorrect answers were incorrect led to deeper understanding of the problems, as indicated by the adoption of strategies that would solve a broader range of problems rather than just the problems in the initial training set.

These findings suggest that explaining incorrect as well as correct answers improves learning. However, the findings may have stemmed from an idiosyncrasy of the data in this study. Unlike the case in numerous other studies, the group that was asked to explain only why correct answers were correct did not learn more than the group that was given only feedback. Thus the greater effectiveness of explaining both correct and incorrect answers than of explaining only correct answers in that study may have been produced by the idiosyncratically low level of performance of the group that explained only correct answers. This concern, as well as the desire to replicate the results and extend them to scientific as well as mathematical reasoning, motivated a study by Siegler and Chen (2008) on children's learning about water displacement.

The task was modeled after Inhelder and Piaget's (1958) water displacement problem. On each trial, first- through fourth-graders were shown two identical beakers containing equal amounts of water. Then the children were shown two objects, told their relative sizes and weights, and informed that either both would float or both would sink. They were then asked which object would cause the water to rise higher if one object were placed in each beaker. The problem's complexity stems from the fact that one variable (weight) determines water displacement when objects float, whereas a different variable (volume) determines displacement when they sink. To state the principle more formally: sunken objects displace a quantity of water equal to their volume; floating objects displace a quantity of water equal to their weight.

This problem was of interest for several reasons. First, it requires differentiation of weight and volume, two quantitative dimensions that are highly correlated in the everyday environment and that even adolescents often confuse (Piaget, 1952). It was also of interest because it addresses the concept of interactions among variables in a particularly direct way. In water displacement, the relevance of all variables depends on the states of other variables. Weight matters when the objects float; volume matters when the objects sink; no variable matters across all displacement problems. A third source of interest was that the task is related to a milestone in the history of science: Archimedes' principle of buoyancy. This principle states that a body immersed in a fluid, either wholly or partially, is buoyed up by a force equal to the weight of the displaced fluid. A floating object displaces an amount of fluid equal to its weight, whereas an object that is totally immersed displaces an amount of fluid equal to its volume (as illustrated in the proverbial tale of Archimedes' insight in the bathtub, in which the mathematician realized that he could determine whether the king's crown was made of pure gold by examining the amount of water it displaced).

The study followed the same type of pretest-training-posttest design as the previously described studies of number conservation and mathematical equality. On each trial in all phases of the study, children were shown two identical transparent glasses filled with water to equivalent points and also a pair of cubes, one for each glass. The cubes varied in weight, volume, and type of material (metal, wood, stone). Depending on the type of material, some cubes floated and others sank. On each trial, children heard descriptions of the cubes' relative sizes and weights (e.g., "this block is bigger than the other one, but they weigh the same") and were encouraged to pick them up. Next, children were told, "Imagine that I put this block into this container and that block into that container, and that both float (or sink). Which container will have a higher water level—or will they be the same?"

Children were randomly assigned to three conditions that paralleled the conditions in the Siegler (2002) study of mathematical equality. In all conditions, immediately after children predicted the effects of putting the blocks in the water, they observed the rise in water levels when the blocks were placed in the glasses and were told by the experimenter "You were right" or "No, that wasn't right."

The three experimental conditions differed in what happened after the children received this feedback. Children in the explain-correct-and-incorrect-answers condition were asked, after the feedback, to explain why the correct answer was correct and then why the answer suggested by the status of the variable that was irrelevant on the trial (weight when the object sank, volume when it floated) was incorrect. The exact question regarding why the correct answer was correct depended on whether the child's prediction on the trial was accurate. If the child's prediction was accurate, the experimenter said: "You were right. Now look carefully at what happened and see if you can figure out why." If the child's prediction was inaccurate, the experimenter said: "No, that's not right. Look carefully and see if you can figure out why that wasn't right. Now tell me why the water level in this container is higher than in that one."

Having explained the correct answer, children in this condition were then presented with the wrong answer and asked to explain why it was wrong. They were told, "A child from another school thought that the water level in this container would be higher than in that one after we put these two blocks into the containers. Why do you think she thought this container would have a higher water level? Do you know why she was wrong?" The experimenter responded with a noncontingent "very good" after the children's explanations in this and the other two conditions.

Children in the explain-correct-answers condition were presented with a procedure that was identical to the first part of the procedure presented to children in the explain-correct-and-incorrect-answers condition. Children in the explain-own-answers condition received only the feedback that children in all groups received; they were not asked to explain the outcome after seeing it.

The number of children who used the correct rule on the pretest (defined as at least 15 of 18 correct answers on the three-choice questions) was zero in all three conditions. The number of children who used the correct rule on the posttest was influenced by both the type of explanatory activity in which they engaged and their age/prior knowledge. Children were divided into an older half (third- and fourth-graders) and a younger half (first- and second-graders); not surprisingly, the pretest knowledge of the older children surpassed that of the younger children. Among the older (and more knowledgeable) children, those who were asked to explain both correct and incorrect answers were more likely to use the correct rule on the posttest than were children who were asked to explain correct answers (67% vs. 37%), and children who were asked to explain correct answers adopted the correct rule more often than children who received only feedback (37% vs. 5%). Among the younger (and less knowledgeable) children, frequency of correct rule use on the posttest showed the same trend (33% vs. 20% vs. 6%), but the differences among the three conditions were not significant. Analyses of the relation between the sophistication of the rules children used on the pretest and their likelihood of learning yielded a similar pattern. Those with the most advanced pretest rules were more likely to use the correct rule on the posttest than those whose pretest rules were less advanced. Age and knowledge on the pretest were sufficiently highly correlated that it was impossible to determine which was the better predictor of pretest performance.

Thus, as predicted by Shrager and Siegler's (1998) and Siegler and Araya's (2005) computer simulations of strategy choice, explaining both why correct answers were correct and why incorrect answers were incorrect resulted in greater learning than only explaining why correct answers were correct, which in turn led to greater learning than receiving feedback but not being asked to explain why the outcome turned out as it did. Stated slightly differently, learning is increased by questions that encourage children to adopt the learning goal of understanding why observed outcomes occur and why other plausible outcomes do not occur.

CONCLUSIONS

The research reviewed in this chapter shows that learning goals cannot be taken for granted. Fifth-grade students in public schools often did not mention even one general learning goal among the five characteristics that they chose for their ideal student. Similarly, students ranging from 5-year-olds to those of college age did not appear to spontaneously adopt the learning goal of explaining their observations regarding number conservation, mathematical equality, and water-displacement problems. If they had, the self-explanation manipulations would have been redundant with the students' spontaneous processing and therefore would have had no effect. After all, the encouragement to explain conveyed no content information whatsoever about any of the three problems—only a suggestion that students adopt the goal of explaining why the observed event occurred.

The research reviewed in this chapter also showed that the adoption of learning goals increases learning. Students who were given information about a teacher's background generated better and more balanced solutions to an interpersonal cross-cultural conflict problem than did students who received only general cultural information. Exposure to the personal information led to superior solutions that took into account and addressed both the teacher's and the students' perspectives. Similarly, students who were encouraged to explain both why correct answers are correct and why incorrect answers are incorrect generated deeper and more general solutions to mathematical equality problems than did students who were encouraged only to explain the correct answers or who received feedback only regarding the correct answer.

Both types of findings argue for both educators and researchers to pay greater attention to learners' goals. Educators cannot assume that students have the same learning goals as they do, nor can they assume that students are trying to explain for themselves observations of physical phenomena, statements of teachers, or passages in textbooks. Although the limited adoption of learning and explanatory goals by many students presents a challenge, it also provides an opportunity; that is, the identification of methods for encouraging students to adopt learning goals seems to have considerable potential for improving education. Further research on learning goals and on ways in which teachers can inculcate them can contribute to this effort.

ACKNOWLEDGMENTS

Support for writing this chapter was provided from grant proposal NSF DRL #0723795 to the first author and the Department of Education IES Grants R305H020060 and R305H050035 to the second

author. The opinions expressed in this chapter are those of the authors only and do not reflect the opinions by the funding agencies listed here.

REFERENCES

Alberts, H. C. (2008). The challenges and opportunities of foreign-born instructors in the classroom. *Journal of Geography in Higher Education, 32*(2), 189–204.

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1- to 2-year-olds' recall of events. *Developmental Psychology, 25,* 197–206.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (4th ed., pp. 77–166). New York: Wiley.

Calin-Jageman, R. J., & Ratner, H. H. (2005). The role of encoding in the self-explanation effect. *Cognition and Instruction, 23,* 523–543.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 1,* 33–81.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development.* Mahwah, NJ: Erlbaum.

Cohen, D. K. (1991). Revolution in one classroom. *American Educator, 15*(2), 16–23, 44–48.

Cole, M. (2002). *Cross-cultural and historical perspectives on the consequences of education.* Paper presented at the Herbert Spencer Lecture, Oxford University, England.

Cuevas, P., Lee, O., Hart, J., & Deaktor, R. (2005). Improving science inquiry with elementary students of diverse backgrounds. *Journal of Research in Science Teaching, 42*(3), 337–357.

Darling-Hammond, L. (1997). School reform at the crossroads: Confronting the central issues of teaching. *Educational Policy, 11*(2), 151–166.

Eini, K. (2005). Friends and flags: A multicultural learning project. *Educational Leadership, 63*(4), 68–70.

Ferguson-Hessler, M. G. M., & de Jong, T. (1990). Studying physics texts: Differences in study processes between good and poor solvers. *Cognition and Instruction, 7,* 41–54.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jacobs, J. K., & Morita, E. (2002). Japanese and American teachers' evaluations of videotaped mathematics lessons. *Journal for Research in Mathematics Education, 33*(3), 154–175.

Jacobs, J. K., Yoshida, M., Stigler, J. W., & Fernandez, C. (1997). Japanese and American teachers' evaluations of mathematics lessons: A new technique for exploring beliefs. *Journal of Mathematical Behavior, 16*(1), 7–24.

Kuhn, D., & Franklin, S. (2006). The second decade: What develops (and how). In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2: Cognition, perception, and language* (6th ed., pp. 953–993). Hoboken, NJ: Wiley.

Lampert, M. (1990). When the problem is not the problem and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27,* 29–64.

Lee, O. (2003). Equity for culturally and linguistically diverse students in science education: A research agenda. *Teachers College Record, 105,* 465–489.

Lee, O. (2004). Teacher change in beliefs and practices in science and literacy instruction with English language learners. *Journal of Research in Science Teaching, 41,* 65–93.

Leslie, A. M. (1982). The perception of causality in infants. *Perception, 11,* 173–186.

Lin, X. D. (2001). Reflective adaptation of a technology artifact: A case study of classroom change. *Cognition & Instruction, 19*(4), 395–440.

Lin, X. D., & Bransford, J. D. (in press). Impact of personal background knowledge on cross-cultural understanding. *Teachers College Record.*

Lin, X. D., & Lehman, J. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research In Science Teaching, 36*(7), 1–22.

Lin, X. D., & Schwartz, D. (2003). Reflection at the crossroad of cultures. *Mind, Culture & Activities, 10*(1), 9–25.

Lin, X. D., & Schwartz, D. (2007). *An ideal student.* Paper presented at the American Educational Research Association, Chicago.

Lin, X. D., Schwartz, D., & Hatano, G. (2005). Toward teachers' adaptive metacognition. *Educational Psychologist, 40*(4), 245–255.

Nathan, M. J., Mertz, K., & Ryan, B. (1994). *Learning through self-explanation of mathematical examples: Effects of cognitive load.* Paper presented at the 1994 Annual Meeting of the American Educational Research Association.

Ngwainmbi, E. K. (2006). The struggles of foreign-born faculty. *Diverse Issues in Education, 23*(10), 28.

Nisbett, R., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review, 108,* 291–310.

Oakes, L. M., & Cohen, L. B. (1995). Infant causal perception. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9). Norwood, NJ: Ablex.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3,* 359–400.

Piaget, J. (1952). *The child's concept of number.* New York: Norton.

Pine, K. J., & Messer, D. J. (2000). The effect of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction, 18,* 35–51.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction, 12,* 235–275.

Prawat, R. S. (1992). Teachers' beliefs about teaching and learning: A constructivist perspective. *American Educational Research Journal, 100*(3), 354–395.

Renkl, A. (2002). Learning from worked-out examples: Instructional explanations supplement self-explanations. *Learning & Instruction, 12,* 529–556.

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70,* 293–315.

Rogoff, B. (2003). *The cultural nature of human development.* New York: Oxford University Press.

Rosenthal, R., & Jacobson, L. (1968). *Pygamlion in the classroom: Teacher expectation and pupils' intellectual development.* New York: Holt, Rinehart & Winston.

Shrager, J. & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9,* 405–410.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28,* 225–273.

Siegler, R. S. (2002). Microgenetic studies of self-explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31–58). New York: Cambridge University Press.

Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed., pp. 464–510). Hoboken, NJ: Wiley.

Siegler, R. S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 33, pp. 1–42). Oxford, UK: Elsevier.

Siegler, R. S., & Chen, Z. (1998). Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology, 36,* 273–310.

Siegler, R. S., & Chen, Z. (2008). Differentiation and integration: Guiding principles for analyzing cognitive change. *Developmental Science, 11,* 433–448.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing Cognitive Competence: New approaches to process modeling* (pp. 31–76). Hillsdale, NJ: Erlbaum.

Staszewski, J. J. (1988). Skilled memory and expert mental calculation. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 71–128). Hillsdale, NJ. Erlbaum.

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 94*(2), 344–355.

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning.* New York: Macmillan.

This page intentionally left blank

5

Working With and Cultivating the Development of Interest, Self-Efficacy, and Self-Regulation

K. ANN RENNINGER

Jane Goodall, the renowned primatologist, remembers climbing into bed with some worms from the garden one night because she wanted to keep on observing them. She also recalls her mother coming in and gently suggesting that the worms be taken back to the garden so that they would be alive in the morning (Goodall, 2008). Jane's situation differs from that of Y, a child of approximately the same age who, on touching a worm, started jumping up and down and screaming that she "broke the family curse" (meaning that neither her mother nor her sister would have touched the squiggly, slimy worm); or does it?

Jane and Y's interest in worms directs their attention; it also impacts the goals that each is ready to set and the forms of instructional conversation, texts, tasks, exhibits, and/or software with which each is prepared to engage. They are different phases of interest, however. Jane is ready to engage in what Yamauchi, Wyatt, and Carroll (2005) term an instructional conversation; she is ready to think about worms and possibly to ask questions about different types of animals and their needs. She is ready to stretch what she presently understands and is being supported by her mother in doing so. Y, on the other hand, may be ready to talk about what she felt when she touched the worm or about either her mother's or her sister's feelings, but she is unlikely to do much in this moment with information about worms and their need to live in soil.

Interest is a cognitive and motivational variable that develops and can be supported to develop (Dewey, 1913); it may reference worms, but it could reference any disciplinary content.[1] As it relates to the process of learning, *interest* refers to the psychological state that accompanies engagement, and it also refers to the likelihood that the learner will voluntarily return to engagement with a particular content of interest (e.g., worms) over time. When learners have a well-developed individual interest, they have relatively higher levels of self-efficacy and are more able to sustain attention, set goals, and use strategies in the discipline of their interest than with content of less interest (Renninger & Hidi, 2002). While they need support from others and the environment to continue to develop their interest (e.g., Barron, 2006; Csikszentmihalyi, Rathunde, & Whalen, 1993), they are also likely to seek and make use of feedback, using it to continue to develop and deepen their interest (see Lipstein & Renninger, 2007a; Renninger & Hidi, 2002). Learners are likely to ask and reflect on their own curiosity questions (e.g., why is it easier to find worms in one location rather than another?)—questions that are not novel to those who have more information but are novel for the learner and allow the learner to build knowledge (Renninger, 2000). Learners' curiosity questions lead them to voluntarily explore and stretch their present understanding, in turn developing ownership and value for knowledge as well as the deepening of interest (Azevedo, 2006; Renninger, 2000; see Flum & Kaplan, 2006).

In contrast, when learners like Y are in earlier phases of interest development for a content, they may be supported to ask curiosity questions and may even experience excitement or pride in the new knowledge they have acquired, but they are not likely to have positive self-efficacy in the sense that they believe they are or can be successful pursuing their interest (especially if they are 8 to 10 years-of age or older; see Renninger, 2009). In earlier phases of interest, they also may or may not be able to self-regulate and are not likely to self-identify with the discipline (Renninger, Bachrach, & Posety, 2008). Instead, they need support to make the kinds of connections that will lead them to develop goals and strategies (Renninger et al., 2008). They need support to inquire, reflect, and make use of the opportunities and/or resources that are available (Renninger, 2009; Renninger & Lipstein, 2006). In earlier phases of interest development, they are primarily dependent on others and the design of opportunities (Lipstein & Renninger, 2007a; Renninger, 2009). Specifics about the forms of support that must be in place to enable a learner like Y to move from heightened affect to the ability to pose, reflect on, and seek answers to curiosity questions about worms are not well understood, however.

Research on motivation has only begun to address the relations among different motivational variables and their relation to changes in learning and motivation over time (Urdan & Schoenfelder, 2006). The continuities that have been identified are rich with implications for how interest might be cultivated and helped to develop through task engagement, although their application can be complicated by differences in the way these terms are used in the vernacular and findings from research. Because terms such as *interest, self-efficacy,* and *self-regulation* have specific meanings in their respective research literatures, this chapter opens with working definitions and an overview of research findings: first of interest and interest development and then of self-efficacy and self-regulation. Following this, case snapshots from several studies are used to identify relationships among the learner's phase of interest and both self-efficacy and self-regulation; the snapshots are then used to discuss implications for practice and future research.

WORKING DEFINITIONS AND RESEARCH FINDINGS

Interest and its development, self-efficacy, and self-regulation are distinct and complementary motivational variables. Their relation appears to change based on the phase of a learner's interest. In the following section, findings from research are used to provide working definitions.

Interest

When interest describes learning that can be cultivated and developed, it refers to both the psychological state of being engaged and to the predisposition to return to engagement over time. This differs from descriptions of rewards or intrinsic motivation (Deci, 1992; Lepper & Henderlong, 2000), vocational pursuit (Armstrong, Allison, & Rounds, 2008; Holland, 1959), and/or positive affect (Alexander, Kulikowich, & Schulze, 1994; Harackiewicz, Barron, Tauer, & Elliot, 2002; Sansone, 1986; Wigfield & Eccles, 2002). Whether a learner needs to be rewarded in order to sustain engagement addresses only one aspect of how sustained engagement unfolds for an individual. While a learner may choose to pursue a vocation suggested by present abilities and skills (vocational interest), such assessment does not account for either the learner's potential to develop new interests over time and/or that a developed interest might regress or disappear without support from others. Finally, positive affect may reflect enjoyment, but it does not account for the capability to set goals, ask curiosity questions, and revise understanding—cognitive characteristics of the development and deepening of interest.

In addition to differing from other conceptualizations of interest, interest as it is discussed here in relation to development also differs from other motivational variables in at least five ways. Interest (a) always refers to engagement with particular content; (b) is composed of stored knowledge, stored value, and feelings; (c) is often an unreflective state or process; (d) develops and is sustained through interaction; and (e) has a neurological basis.[2]

First, interest always refers to a learner's engagement with a particular content—for example, worms. Other motivational variables more typically reference an individual's characteristics for a specific task or across the range of his or her engagements. Thus, a learner might be described as having the ability or inability to self-regulate.[3]

Second, interest includes three components: knowledge, value, and affect; these exist and develop in relation to the learner's other engagements. In early phases of interest development, knowledge may be minimal; as interest develops, knowledge supports and contributes to both value and feelings. Jane's interest in worms, for example, varies from the phase of her interest for other activities, such as practicing the piano, reading, and so on. Her existing knowledge and values presumably led her to want to continue to observe the worms even though she had to go to bed. The supportive and informative nature of her mother's response probably encouraged her developing interest. It also provided her with the information that the worms need soil to live—knowledge that is associated with the value she already held for observing worms. Other motivational variables do not typically reference or assess learner knowledge about content but instead are conceptualized as feelings that influence beliefs about success, competence in setting goals, and who the learner is.

Third, interest is often an unreflective state or process. A learner's interest may be triggered without the learner being aware of the triggering process. Jane's description of her mother's support of her interest in the worms is hindsight; her interest in worms triggered her idea that she could continue to observe the worms in bed and then again the information that the worms might die maintained her interest. In neither case did Jane focus on the fact that her interest was triggered; rather, she experienced the triggering and could later talk about it. Similarly, Y's interest was triggered when she touched the worm. There was no immediate support for her to ask a question or reflect on her experience, and according to the observer notes, she moved on soon after to do some dance steps with a friend; her triggered interest was not maintained. Interest is not something learners (especially young learners) can typically will themselves to experience; this is different from other motivational variables associated with volition (Eccles, Wigfield, & Schiefele, 1998). However, with age, learners are increasingly able to self-regulate to develop interest (Sansone, Weir, Harpster, & Morgan, 1992) and be self-determined (Ryan & Deci, 2000).

Fourth, interest develops through interactions with other people and the nature of the tasks and opportunities of the environment (see also Barron, 2006). Jane, for example, had an interest in observing worms, and her interest was both supported and stretched in her mother's interaction with her. If her mother had responded to the worms in horror and/or had been furious about the dirty bed sheets, the development of Jane's interest might have been different. Interest is promoted through interactions with the environment, whether these interactions are initiated by others or by the learner. Although interest is often described in relation to a person, it is more accurately described as an interaction of the person and the environment. In this sense, then, it is always both extrinsically and intrinsically supported; it is not simply a form of intrinsic motivation (see discussion in Hidi, 2000). In contrast, motivational variables are typically described as characteristics of the person and are often operationalized as binaries in the research literature: the learner has (or does not have) feelings of self-efficacy and is (or is not) self-regulated.

Fifth, interest has a physiological basis. Brain reactions can be expected to differ when a learner is and is not engaged with an identified content of interest (Ashby, Iseni, & Turken, 1999; Hidi & Ainley, 2008). In the neuroscientific literature, interest-based activities are referred to as "seeking behavior," the developing capacity of learners to pursue answers to their own curiosity questions (Panksepp, 1998; see discussion in Hidi, 2006). The information that is sought depends on the biological, psychological, social, and physical capacities of the learner, including the learner's ability to perceive and make use of available opportunities to engage content (Renninger & Lipstein, 2006). Because seeking behavior has an identified physiological basis, differences among learners in terms of interest may exist in the particular context and/or the phase of interest but not in the process and role of interest in learning and development (Renninger, 1990). Both self-efficacy and self-regulation, on the other hand, do change in relation to cultural norms and expectations (see, e.g., the discussion in Roeser, Peck, & Nasir, 2006).

Development of Interest

A learner can be expected to be in one of four phases of interest development with respect to any given disciplinary content (Hidi & Renninger, 2006); (see Table 5.1). The phases are sequential but are not invariant—meaning that interest may develop, but it may also regress or fall off (Bergin, 1999; Renninger & Lipstein, 2006). Interest is dependent on what the learner perceives, how she or he represents or connects this to prior understandings, and how this informs his or her activity (see Renninger, 1990, 2000). An educator could intend to provide opportunities for interest to develop (e.g., group work), but the effectiveness of this is related to how and whether these opportunities are recognized and seized by the learner. Learners in each phase of interest development are differently positioned to work with opportunity.

In the first phase, *triggered situational interest,* a learner might attend briefly to a worm. Interest could be triggered by the worm's slimy and squiggly features and maybe by others' disgust. Examples of other triggers for situational interest include the glitz of software design (Lepper & Cordova, 1992; Cordova & Lepper, 1996), the panic induced by thinking that a project had failed (Renninger & Hidi, 2002), or familiar content (e.g., bicycles, musical instruments) used as a basis for scientific explanation (Hoffmann, 2002). Triggers draw attention to particular content, but, the learner has so little stored knowledge or stored value for the content in the earliest phase of interest development that he or she is not likely to pursue further work without external support. Principled knowledge and the sense of competence that emerges with valuing (Renninger, 1990, 2000; White, 1959) appear necessary for the learner to make a connection to content.

Regardless of how many or which well-developed interests a learner has, the development of a new interest begins with a triggered situational interest. Once the interest of a learner is triggered, it may or may not develop into a *maintained situational interest.* Understanding what enables a shift from a triggered situational to a maintained situational interest is an open question for interest research; it appears that such shifts are related to support(s) from the environment in one or another

Table 5.1

LEARNER CHARACTERISTICS, FEEDBACK WANTS, AND FEEDBACK NEEDS IN EACH OF THE FOUR PHASES OF INTEREST DEVELOPMENT

	PHASES OF INTEREST DEVELOPMENT			
	PHASE 1 – TRIGGERED SITUATIONAL INTEREST	**PHASE 2 – MAINTAINED SITUATIONAL INTEREST**	**PHASE 3 – EMERGING INDIVIDUAL INTEREST**	**PHASE 4 – WELL-DEVELOPED INDIVIDUAL INTEREST**
Learner Characteristics	Learners: ■ Attend to content, if only fleetingly ■ Need support to engage: – From others (e.g., group work, instructional conversation) – Through instructional design (e.g., software) ■ May experience either positive or negative feelings ■ May or may not be reflectively aware of the experience	Learners: ■ Reengage content that previously triggered attention ■ Are supported by others to find connections between their skills, knowledge, and prior experience ■ Have positive feelings ■ Are developing knowledge of the content ■ Are developing a sense of the content's value	Learners: ■ Are likely to independently re-engage content ■ Have curiosity questions that lead them to seek answers ■ Have positive feelings ■ Have stored knowledge and stored value ■ Are very focused on their own questions ■ May have little value for the canon of the discipline and most feedback	Learners: ■ Independently reengage content ■ Have curiosity questions ■ Self-regulate easily to reframe questions and seek answers ■ Have positive feelings ■ Can persevere through frustration and challenge in order to meet goals ■ Recognize others' contributions to the discipline ■ Actively seek feedback

(*Continued*)

LEARNER CHARACTERISTICS, FEEDBACK WANTS, AND FEEDBACK NEEDS IN EACH OF THE FOUR PHASES OF INTEREST DEVELOPMENT *(Continued)*

	PHASES OF INTEREST DEVELOPMENT			
	PHASE 1 – TRIGGERED SITUATIONAL INTEREST	PHASE 2 – MAINTAINED SITUATIONAL INTEREST	PHASE 3 – EMERGING INDIVIDUAL INTEREST	PHASE 4 – WELL-DEVELOPED INDIVIDUAL INTEREST
Feedback Wants	Learners want: ■ To have their ideas respected ■ Others to understand how hard work with this content is ■ To simply be told how to complete assigned tasks in as few steps as possible	Learners want: ■ To have their ideas respected ■ Concrete suggestions ■ To be told what to do	Learners want: ■ To have their ideas respected ■ To express their ideas ■ *Not* to be told to revise present efforts	Learners want: ■ To have their ideas respected ■ Information and feedback ■ To balance their personal standards with more widely accepted standards in the discipline
Feedback Needs	Learners need: ■ To feel genuinely appreciated for the efforts they have made ■ A limited number of concrete suggestions	Learners need: ■ To feel genuinely appreciated for the efforts they have made ■ Support to explore their own ideas	Learners need: ■ To feel that their ideas and goals are understood ■ To feel genuinely appreciated for their efforts ■ Feedback that enables them to see how their goals can be more effectively met	Learners need: ■ To feel that their ideas have been heard and understood ■ Constructive feedback ■ Challenge

form (Alexander, 2004; Schraw & Lehman, 2001; Silvia, 2006). Group work, computers, games, and the meaningfulness of tasks have all been identified as methods that lead learners to reengage (Mitchell, 1993; see also Palmer, 2009). While learners with a maintained situational interest typically have positive feelings and may reengage due to support, they may not yet have enough knowledge and value for the content to be able to identify goals and confidently pursue their own question(s) independently (Renninger et al., 2008). They often want to be told what to do and how to do it (Lipstein & Renninger, 2007a, 2007b). At the same time, learners in this phase of interest development need to ask questions, explore, and begin to take ownership of the content if they are going to continue to develop and deepen their interest (Flum & Kaplan, 2006; Renninger, 2009; Renninger & Lipstein, 2006).

For the learner, there is a tension between the desire to be told what to do and the need to take a chance and explore core ideas of the content (Renninger & Lipstein, 2006; see also Lipstein & Renninger, 2007a) that may account for the difficulty researchers have had describing the transition from a maintained situational interest to the next phase, an emerging individual interest (see Schraw & Lehman, 2001). Krapp (2002) describes the phase of maintained situational interest as "working interest," suggesting that it only "works" if there is external support for the learner and the possibility of interest development. In actuality, of course, the learner's need for support continues in later phases of interest, but the forms of support that need to be in place shift from being primarily external (other people enabling connections and engagement) to being primarily internal (the learner identifying connections and seeking engagement) (Hidi & Renninger, 2006). Learners in each phase of interest do need external support, however. In earlier phases of interest they may need to have their success pointed out to them and their competence acknowledged, whereas learners in later phases of interest have some sense of their capacities relative to the discipline and need interactions in which they are taken seriously and their competence is acknowledged. Jane carried the worms to bed because she considered this to be a reasonable place to continue her observations. She devised a way to continue her observations, and her mother encouraged and informed her with her response. Y, in contrast, would probably need support from others to observe worms.

Retrospective interviews indicate that learners can move rather quickly through the phase of maintained situational interest and on to the third phase, an *emerging individual interest* (Lipstein & Renninger, 2007a). One boy, for example, described reading a novel he wrote to his elementary school classmates, being hailed as an author, and knowing that he was a writer. Once he was acknowledged as a writer and assumed this identity, he began engaging and exploring writing in a different way (Renninger & Lipstein, 2006). In other examples, an existing interest can trigger interest in a new domain. Thus, a teacher with a well-developed individual interest for math had her interest triggered and maintained by the possibilities for further understanding math through developing knowledge of technology (Renninger & Shumar, 2002). The possibilities for this teacher include both developing knowledge of technology and recognizing its utility. Similarly, a teacher with a well-developed interest in math had his interest in pedagogy triggered and maintained by his collaborations with other teachers who had a well-developed interest in math pedagogy (Renninger & Shumar, 2004). In each case, the coupling of interest for different content provided a context that enabled a different way of engaging new content, which, in turn, led to sustained engagement and exploration (Krapp & Fink, 1992).

Learners who have an emerging individual interest generate curiosity questions based on their knowledge of the discipline. Curiosity questions appear to enable the learner to shift from what could be considered primarily extrinsic to primarily intrinsic involvement with content. In this phase of interest, learners are excited by their new understanding of content and are focused on their own questions. As a result, they are not always receptive to feedback (Lipstein & Renninger, 2007a). Rather, learners in this phase of interest development appear to relish autonomy and assert it (Azevedo, 2006; Barron, 2006). They can seem disrespectful and disinterested because they are so deeply invested in developing their own understanding that they distance themselves from others' ideas (Lipstein & Renninger, 2007a).

The continued development of interest into the fourth phase, *well-developed individual interest,* appears to occur when the learner is able to sustain the forms of self-regulation necessary to pose and reframe curiosity questions based on available feedback and resources (Barron, 2006; Lipstein & Renninger, 2007a; Renninger, 2000). For learners with a well-developed individual interest, the overall experience of working with a content of interest is positive. Presumably because of this and their principled knowledge of the domain, they have a kind of long-range vision of what steady work can yield. They also have enough confidence to both persevere when they run into difficulty (Prenzel, 1992; Renninger, 1992; Renninger & Hidi, 2002) and to seek and work with feedback that involves revising what they thought they knew (Lipstein & Renninger, 2007a). Learners in this phase of interest also readily seek information that allows them to balance their personal standards for working with content with more widely accepted standards in the field (Lipstein & Renninger, 2007a). Importantly, while they provide their own reinforcement for reengaging the content of interest, others also acknowledge them for the depth and quality of their work (Hidi & Harackiewicz, 2000).

Learners with a well-developed individual interest are ready to work with and create opportunities for themselves and, as a result, seek external support in the form of feedback from others. They are also likely to have high levels of self-efficacy and self-regulation and to identify with others who seriously pursue their content of interest, although dips in their interest, self-efficacy, and/or their self-regulation can occur given changes in context (see Eccles & Midgley, 1989; Marsh, 1991). In earlier phases of interest development, the phase of learner interest does not so readily map onto a particular level of self-efficacy or self-regulation. In these earlier phases of interest, learners need support to recognize and engage opportunities; the extent to which they feel and respond to support impacts their future engagement and view of themselves in relation to the content. More explicit consideration of the relation among the phase of learner interest and both self-efficacy and self-regulation may well offer insight into the kind of support that learners in earlier phases of interest need.

Self-Efficacy

Self-efficacy refers to a learner's beliefs about his or her ability to succeed at tasks (Bandura, 1986, 1997)[4] and is more typically linked to judgments of capacity to engage a task than to the skills the task requires (Zimmerman & Schunk, 2004). Positive self-efficacy is generally considered to be a predictor of success because it has been found to influence strategy use, effort, and perseverance (e.g., Bouffard-Bouchard, Parent, & Larivee, 1991; Lent, Lopez, & Bieschke, 1991; Pajares, 1996; Pintrich, 2004). While general ability has been found to influence the accuracy of learner perceptions about success (e.g., Pajares & Kranzler, 1995; Zimmerman, Bandura, & Martinez-Pons, 1992), age is also a factor. Younger children are likely to have more positive perceptions of their abilities to achieve than are older children (see Harter, 1985, 1986; Wigfield & Harold, 1992) and are, as a result, more likely to persevere, especially if they feel supported.

With the increasing ability to assess and compare others' abilities to their own, children's perceptions of their own abilities typically decline (Harter, 1999, 2003); in classrooms particularly, comparison to peers and teacher evaluation affects self-efficacy and performance (Eccles & Midgley, 1989). The presence of environmental supports such as positive feedback (Bouffard-Bouchard, 1990; Cervone & Peake, 1986; Eccles & Midgely, 1989) and teachers and parents who communicate that the learner can be successful can positively affect learners' sense of their own possibilities, however (Eccles & Midgley, 1989; Eccles, Wigfield, Harold, & Blumenfeld, 1993; Nieswandt, 2007; see also Yamauchi et al., 2005).

Like interest, then, self-efficacy is characterized by the feelings and valuing that accompany competence. Unlike interest, feelings of self-efficacy are not defined as developing in relation to principled knowledge; it is possible to have beliefs about possibility that are not founded in knowledge.

Unlike self-efficacy, interest is not a belief; although learners may hold beliefs about content that is of interest to them. For example, they may believe in the idea that physics is hard (see Hannover & Kessels, 2004; Kessels, 2005) and this belief may affect their feelings of efficacy and the trigger(s) they then need for interest.

Self-Regulation

Self-regulation refers to a learner's ability to self-structure his or her activity in order to attain his or her goals. As Linnenbrink and Pintrich (2000) suggest, these goals may take one or more forms. They may include goals to complete an activity or task (Harackiewicz & Sansone, 1991; Schunk & Zimmerman, 1997); life goals (e.g., for superiority or happiness, Ford, 1992); and achievement goals (learning and performance goals, Dweck & Leggett, 1988; task goals and performance goals, Maehr & Midgley, 1991); or mastery or performance goals, Elliot & Harackiewicz, 1996).

Research suggests that the more achievement-oriented the goals of the learner are, the more effective his or her self-regulation on school tasks, on the playing field, and in out-of-school activities such as darts and piano playing will be (Hidi & Boscolo, 2006; Kitsantas & Zimmerman, 2002; Pintrich, 2004; Pintrich & Zusho, 2002; Zimmerman & Bandura, 1994; Zimmerman & Kitsantas, 1997, 2006; see also Hidi & Ainley, 2008; Schunk & Zimmerman, 2008). Research findings also distinguish between learners' self-set goals and assigned or expected performance goals as well as between achievement goals and mastery; they indicate that self-set goals are optimal (Harackiewicz, Barron, Carter, Lehto, & Elliot, 1997; Harackiewicz & Elliot, 1993; see also Barron & Harackiewicz, 2000). If learners can sustain their own goals for a task, they develop value and interest; however, if they encounter obstacles to their goals, they must either invoke strategies that allow them to persevere or modify their original goals (Boekaerts, 2006).

The ability to set goals and follow through to achieve them characterizes learners in the last years of high school and beyond, as well as younger learners who have a well-developed interest (Renninger, Sansone, & Smith, 2004). Presumably because they can weigh and rationalize competing goals (e.g., a dislike of organic chemistry yet a desire to do well in order to qualify for medical school), older learners are in a position to identify obstacles and self-regulate (see Husman & Lens, 1999; Pintrich, 2004; Sansone & Smith, 2000). The capacity to self-regulate during task engagement is typical of younger individuals only if they have a developed interest for content to be learned.

In fact, younger learners with a well-developed individual interest for content self-regulate easily and may even do so without the intent to self-regulate (see Renninger & Hidi, 2002). They have goals and they want to accomplish them, but this does not appear to require the reflective awareness that they have a goal; instead, their goals and their actions are coordinated (Renninger, 1990, 2009). Thus, a 13-year-old with a well-developed individual interest for soccer may seem to be always kicking a ball and positioning himself to find opportunities to play, whereas the same learner working with a content for which he is in an earlier phase of interest development (e.g., science) is likely to need support (e.g., set times for and monitoring of work on a science project, clear goals for project completion) to begin to self-regulate (see Renninger & Hidi, 2002). Other learners with less developed interest for soccer need support for self-regulating to play soccer (e.g., they are scheduled by others for practices, must be reminded to practice, have to have the steps of a drill specified and repeated); these learners may or may not be able to self-regulate to engage science content.

Younger learners and those in later phases of interest development are not as effective at self-regulating, however, if the learning environment requires them to accommodate the demands of others who have ideas that vary from their own. Because they either have or think they have both knowledge and value, they may not be able to work with others' ideas when they differ from theirs. Thus the soccer player with a developed interest in soccer who is made to do drills has difficulty self-regulating if, from his perspective, these do not add to his capacity. He may even sacrifice the opportunity to play rather than reevaluating his goals (and attitude). Likewise the learner with a well-developed interest for mathematics who is ready for and values working to understand alternate solution paths may find it difficult to remain engaged in the class of a math teacher who wants to cover a lot of content (see Renninger, 2000).

Like interest, then, self-regulation is a process of cognitive and affective engagement. Unlike interest, self-regulation involves the strategies that a person does (or does not) invoke in order to attain his or her goals. Of note, when the learner has a well-developed interest, these strategies may be identifiable by others but may also be so well-integrated or automatic that the learner is not likely to be fully conscious of them.

CASE SNAPSHOTS, QUESTIONS, AND IMPLICATIONS

In a few studies that employed qualitative methods (Azevedo, 2006; Barron, 2006; Gisbert, 1998; Nolen 2006; Pressick Kilborn & Walker, 2002; Renninger et al., 2008; Renninger & Hidi, 2002), the role of exploration is highlighted and the learner's feelings of self-efficacy and skills in self-regulating appear to have a reciprocal relation to interest development. Quantitative research on interest, however, has focused on learners in one or another phase of interest development. With the exception of the work of Eccles and colleagues (e.g., Denissen, Zarrett, & Eccles, 2007), whose measures assess earlier phases of interest development, quantitative research has not explicitly explored the relation among interest, self-efficacy, and/or self-regulation. Clarification of changes that occur in the transition from one to another phase of interest with respect to the interplay of these variables is needed if research is to inform understanding of how interest might be worked with and cultivated.

As an interim step in this direction, case snapshots from research on interest development and questions that might be asked in practice are presented and discussed, using data from the research on which the snapshots draw.[5] The snapshots span a variety of disciplinary content in order to (a) identify the relations among the phase of a learner's interest and the variables of self-efficacy and self-regulation, (b) highlight the generalizability of the structure and content of each phase in the development of interest across content, and (c) allow ease of referencing throughout the rest of the chapter.

Snapshot 1: Triggered Interest—Dissecting Worms

> If you saw a group of 9- to 12-year-old girls and boys jumping up and down and squealing delightedly while participating in a worm dissection during your summer Biology Workshop, would you think that, because of their heightened affect, they were interested and therefore motivated learners? Could you imagine that they might identify themselves as potential scientists? What if you learned that they were inner-city children who came from a community where there were no adult role models who were scientists and that if their school did have science classes, they offered only facts about science and focused on the memorization of facts? If you were running the Biology Workshop, would you offer these children choices about what to do? Would you still think that they were interested, motivated, and likely to pursue careers in science? (This snapshot is based on studies reported in Renninger et al., 2008)

Just because the learners in this snapshot are excited, we cannot assume that they have a developed interest in science, positive self-efficacy, or the capacity to self-regulate. In order to draw reliable conclusions about their phase of interest, we would need to know that they had enough knowledge to set goals for themselves as learners and that they could independently return to work with this type of content without support (see Renninger et al., 2008). The fact that they were all working together as a group could suggest that they might independently return to talking and thinking about what they were learning and that their interest for the group might help them to make connections to the science content even with little prior knowledge (Sansone & Smith, 2000). Findings from pre- and postworkshop interviews and tasks, however, indicate that although the children liked the workshop a lot and described science as fun, they (a) did not yet have a principled knowledge of either the workshop topics or science more generally, (b) considered science to be hard, and (c) did not imagine that they would ever choose to pursue science.

The children's prior experience and consequently their lack of stored knowledge and stored value constrain their readiness to think about biological principles such as relations among form and function. Information about both their background and what they understood prior to the workshop is especially important for those developing the workshop and facilitating it. Armed with such information, they could adjust examples and expectations so that any interactions start with what the learners appear to be noticing—for example, providing challenges that are optimal and not overwhelming (Pea, 2004; Vygotsky, 1978). Such information is also essential to understanding the beliefs and habits that are already in place (Kilpatrick, Swafford, & Findell, 2002) and for supporting the children to know that they are learning science.

If the school exposure to science that they do get is predominantly about facts and memorization, children are not likely to ask questions in the Biology Workshop without support, nor are they likely to understand that questioning is a fundamental component of science. Similarly, if no one in the children's home or school environments is a scientist or engages them in thinking about science, they are not likely to have a counternarrative about science as a possibility for themselves (Renninger, 2009; see also Perry, 2003). Moreover, research on choice further suggests that children in this phase of interest need minimal guidance (Flowerday & Schraw, 2003). Choice requires more information than these children presently have; they need to be able to appreciate differences among the options in order to effectively make a choice; without the content knowledge and accompanying valuing that characterize more developed interest, learners simply pick at random rather than making a choice (Katz & Assor, 2007).

Snapshot 2: Maintained Situational Interest—Latin Class

> Working-class students in a ninth-grade Latin class all signed up for a second year of Latin study with the exception of one person who dropped out of school. The students come to class each day with a wide range of "Latin Moments": accounts of the instances since the last time the class met in which they encountered a Latin word or phrase, references to historical figures, architecture, or events in soap operas, movies, other classes, etc. They are engaged and attentive; in fact, the Latin program in the school is expanding because so many students are choosing to continue to study Latin. Would you say that the students in this class have an interest in learning Latin? Are these students developing identities as students of Latin? (This snapshot is drawn from data presented in Renninger et al., 2004)

With the exception of the student who dropped out of school, the students in this Latin class appear to have more knowledge and value for learning Latin than the children in the Biology Workshop had for learning about biology. They have elected to pursue a second year of Latin (see Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000; Harackiewicz et al., 2002), describe the class as fun, and know that they are learning Latin. The assignment to identify and report on the Latin Moments they encounter between class periods keeps them engaged. The design of the activity helps them to self-regulate by giving them a clear set of goals, providing models of how to be successful reporters of their own activities, and engaging them in an activity that feels meaningful (see Mitchell, 1993; Palmer, 2009).

It is possible to imagine, based on their response, that the students feel that they can do this activity—that their feelings of self-efficacy are positive (Bandura, 1986). It is not clear, however, whether they are working on developing a principled knowledge of Latin, or if, like the worm dissection, the activity simply draws them in and may have given them some basic knowledge. There is no evidence of questioning, and it is possible that they are not yet ready for more than the kind of guided exploration that the Latin Moments activity offers.

The activity involves them in telling others about their own lives and hearing about others' lives; it also involves them in identifying what could be connections between their own lives and ancient Rome and Greece and the Latin language (e.g., the name of Optimus Prime, on the TV show/movie titled *The Transformers*). An open question is how the activity might be modified in order to fade or lessen ongoing external support for self-regulating (Pea, 2004), in turn enabling them to pose and seek answers to their own questions about the connections they are making. Inasmuch as the assignment scaffolds the students to make connections to Latin as a language and culture, it presently constrains their possibilities. They are supported to notice the occurrence of Latin and to make choices, but the activity does not extend beyond identifying and reporting out.

Like the children in the Biology Workshop, these students have no models other than their teacher and their classmates as possible classicists. Because of this, they may need more than one point of connection to the content in order to undertake problem solving in the discipline of Latin and to consider its possibilities for themselves. They also need opportunities to think and talk about what they notice and wonder if their interest is going to continue to develop and deepen (see examples of noticing and wondering in mathematics in the activity series at http://mathforum.org/pow/support/activityseries/geopow_psc.400.pdf; http://mathforum.org/pow/teacher/PoWsGettingStarted.pdf).

Working with what the students of Latin observe and know, it should be possible to help them stretch their knowledge, either by wondering out loud and modeling ways to make use of existing resources to find answers or by pointing out parallels between what they think and some basic information. For example, the students might be asked why they think that Optimus Prime is called that and asked why they think so, and so forth. The first part of the name Optimus means "best" in Latin. It can be a masculine adjective or a noun meaning "best, best of men" and is most often used to describe certain people in Roman government, especially nobles and aristocrats. The Optimates were a group of people who supported the aristocrats during the Roman Republic. It is the superlative of the word *good,* which is *bonus* in Latin. What the students might think that they want and like—the way that their class now operates—and what they need if they are to make the transition to a more developed phase of interest and more independent learning appears to be content that will stretch what they presently know (see Table 5.1).

Snapshot 3: Emerging Individual Interest—Middle-School Writers

> Imagine that you are teaching middle-school students English. Most of the students do not read the feedback that you give them on their writing. A few of your students act as though they knew how to write and are almost rude about the suggestion that you might teach them anything—saying things like your class is too easy and does not challenge them enough. Your assessment is that these students are not strong writers: the organization of their essays is weak; they write voluminously and use big words unnecessarily; and they do things that drive you wild, like including extra exclamation points in their compositions. Would you consider them to be interested and motivated learners? Would you offer this type of student a choice about either the content or genre of an assignment? Would you expect that such a student might already or even eventually identify as a writer? (This snapshot is drawn from portraits reported in Lipstein & Renninger, 2007a)

The students in this phase of interest are motivated to work on their writing. They have positive feelings about writing and have both stored knowledge and stored value. They do not necessarily fit the textbook case of the interested and motivated learner, however, because they are still developing their skills. They do not set goals that look like those that others (e.g., teachers) might set for them. As a result, they are not self-regulated in ways that others might wish, and their feelings of self-efficacy do not predict achievement as the literature currently suggests (Linnenbrink & Pintrich, 2000). For example, learners in this phase of interest want to hear that their ideas are good and often do not want suggestions for revising their writing. Unlike students in earlier phases of interest who want to be told what to do, students in this phase of interest want to be independent and recognized for the work that they are doing.

Students in this phase of interest have curiosity questions in the sense that they have ideas about how to express themselves. They are exploring expression by using words that have multiple syllables, analyses that hold meaning for them even if others cannot parse them, and punctuation that may not be conventional. They have begun to identify as writers and want to succeed; they can also appear to be disrespectful and not interested.

Snapshot 4: Well-Developed Individual Interest—the Cellist

> What if you were in charge of assigning students to music ensembles at a college with only a few excellent musicians and one day a promising cellist requests assignment to a group with other strong musicians. Is she in a position to make choices about her needs as a learner? Would you assign her to an ensemble with other strong musicians? Would you choose instead to distribute the expertise across ensembles and assign her to a group that included less strong musicians? How would each of these decisions impact her interest and motivation to continue to play the cello in ensemble? (Like the other snapshots, this is an authentic case).

The cellist is promising because her studies with a member of a symphony orchestra have led her to relearn some of her fingering, and this has led her to ask questions about interpretation that she had not asked before. As a learner with a well-developed individual interest, the cellist has positive self-efficacy, clear goals for herself, and the ability to self-regulate. She practices long hours with ease, and she identifies herself as a musician. Given the constraints of the program described, the cellist could of course be encouraged to think about her desires in relation to the group. For her to refocus her interest to the group, she would need to shift her focus from playing and continuing to develop her capacity as a musician to helping others to develop as musicians. As a person with a well-developed individual interest, she feels that she is in a position to make informed choices about what she needs even if it conflicts with what others would like or even recommend. Since the time she would spend in the ensemble is the only time that she has to play the cello, it is not surprising that after she was assigned to the ensemble with weaker musicians, she elected to find other opportunities to play the cello outside of the music program.

For interest to be sustained, learners with well-developed interest, like those with less-developed interest, need to continue to deepen and develop their understanding of content—their interest is fueled by opportunities to engage, explore, and be challenged. Simply being able to play the cello, for example, is not likely to be sufficient. Of course, the cellist could have identified some reason to decide to participate in the less rigorous ensemble group (e.g., the opportunity to spend time with others who are friends, or challenges such as sight-reading a genre with which she is not familiar, or teaching others) and then chosen to do so. In the case described, she did not.

IMPLICATIONS FOR PRACTICE AND FUTURE RESEARCH

What is known then about how to work with and cultivate the development of interest in addition to self-efficacy and self-regulation? The preceding sections of this chapter indicate that learners in each phase of interest development are likely to have distinct motivational profiles. They also indicate that learners in each phase of interest development need support to cultivate and develop interest. However, while all learners may need support, the forms of support provided should differ based on the learner's phase of interest, feelings of self-efficacy, and ability to self-regulate. In other words, responding to a learner such as Y, in an earlier phase of interest development, might ideally involve helping her to clarify what she understands and has experienced, so that this new understanding can be used in some way. In contrast, support for a learner with more developed interest, like Jane, might involve providing additional information in the context of present activity, creating opportunities to think together about what is noticed, and/or modeling of alternate lines of reasoning.

In earlier phases of interest development, interest is in the process of being developed, and so are learners' feelings of self-efficacy and abilities to self-regulate. In these phases, learners do not have a lot of principled knowledge about content. As depicted in Table 5.1, they want to be respected, and they need support in order to know how to engage tasks and begin to develop an understanding that there are multiple ways to work with such tasks successfully.

Y, the children in the Biology Workshop, and the learners in the Latin class all need support from others if they are to either begin or continue to make the kinds of connections to content that will set them up to ask curiosity questions and become independent learners of that content. Such connections might be enabled by either instructional conversations in which the learners' thoughts are respected and become the basis of discussion (Yamauchi et al., 2005) or the design of tasks that encourage them to ask questions and seek information because they want to and not because they have to (Quintana et al., 2004; Zahorik, 1996; see also Sansone & Smith, 2000).

An exercise such as Latin Moments may enable the students to make connections, but because it simply leads to connections rather than also engaging them in wondering and noticing the types of connections they are making, it seems likely that the phase of most students' interest will not develop further. The students have positive self-efficacy but they do not need to self-regulate to do more than find examples of Latin Moments. The activity triggers but does not appear to sustain their interest because, like the children in the Biology Workshop, they may not yet have enough knowledge to ask questions and begin seeking their own questions. It seems that they need to be supported to ask questions and find answers to questions that they pose themselves. The questions and assignments that the teacher poses for them do not necessarily provide this kind of support.

As depicted in Table 5.1, the students in the Latin class, at least initially, may prefer to receive the clear directive to identify Latin Moments rather than to explore the meaning and relation among the moments and the content of Latin class. In another study of a different cohort of children in the Biology Workshop, an ICAN intervention was coupled with the triggers of interest that were already part of the Workshop (Renninger, Bachrach, Riley, Niwagaba, & Tibbets, 2009). The intervention consists of the children writing reflections, ICAN statements, about their learning of the day's instructional objectives. It was hypothesized that this coupling would enable the children to know both that they were learning science and that they could learn science. Findings from this study indicate that over the 5 weeks of the Workshop, the children's self-efficacy for science increased, their interest for science developed, and they outperformed a control group in terms of the types of questions that they began to ask about science and the strategies they employed.

Coupling the ICAN intervention with triggers for interest appears to enable the children to know that they are developing an understanding of science. Presumably this, in turn, also supports them to begin asking questions, using resources to find answers, feeling that they (and not others) are in control of their learning, and that they have the possibility of "doing" science. These findings complement those of Nieswandt (2007), who reports that the presence of both situational interest and self-concept of ability contribute to the development of conceptual understanding.

Learners whose phase of interest is an emerging individual interest, in contrast, can begin to take control of their learning by refusing to work with feedback; they seem to know that they have the possibility of seriously engaging their content of interest but may not yet recognize that feedback could make a contribution to this process (see data-based presentation in Lipstein & Renninger, 2007a). Despite positive self-efficacy, they either do not choose to ask for or work with feedback, or they may not have enough self-efficacy to ask for and receive feedback. While they self-regulate, sometimes writing for hours, their self-regulation is directed to goals that they have set for themselves and not necessarily to the assignments they have been given. It seems that they are not yet able to also give attention to feedback or others' ideas.

Learners with a well-developed individual interest for content to be learned, on the other hand, typically appear to have higher levels of self-efficacy and the capacity to self-regulate. They are clear about their goals and feel enough self-efficacy to recognize that they need feedback and challenge in order to sustain and develop their interest. As a result, they are positioned to work with feedback and others' ideas. Like the cellist, they are also clear about their needs. While the cellist as an older learner is in a position to make decisions about whether to participate in an activity that may not be challenging in the ways she wishes, the younger learner often does not have this option. When children are not able to get to distant athletic practices with a more challenging league, or are discouraged from identifying alternate solution paths in math, and/or are provided with information they already know, the development of their interest is constrained and is likely to fall off and be replaced by other interests (Renninger, 2000).

Aside from the increased likelihood of self-comparison beginning around 8 to 10 years of age, the process of working with and cultivating interest appears to be driven more by experience than by age-related considerations—meaning that learners at any age can have more developed interests and also that they can be supported to make connections to and develop new interests.

In earlier phases of interest development, learners may click randomly while working online or be taken by the glitz of a program's design if noticing and wondering is not encouraged. Learners in these phases of interest development typically need support to know what kinds of goals are possible and also how these might be realized (Renninger et al., 2008). It probably is not the case that these can simply be explained, although they may be modeled (see Quintana et al., 2004; Schwartz & Bransford, 1998).

Thus, working with and cultivating the interest of learners in earlier phases of interest development involves being informed about both their motivational profiles and their readiness to engage disciplinary content. Their motivational profiles includes their (a) feelings of self-efficacy, and (b) readiness to self-regulate in working with these tasks. Their readiness to engage disciplinary content includes the presence of (a) triggers to attend to new content; (b) guidance to support them to make connections to this content (Kirschner, Sweller, & Clark, 2006; Yamauchi et al., 2005; Zahorik, 1996); and (c) requirements like those of the ICAN statements, which encourage learners to make sense of their work, manage it, and articulate and reflect on their connections (Quintana et al., 2004)—capacities that characterize learners in later phases of interest.

Working with learners in more developed phases of interest also requires consideration of their motivational profiles and their engagement with disciplinary content. However, learners in later phases of interest development need less direction to reflect and more opportunities to continue to stretch their knowledge and engage in what Azevedo (2006) terms personal excursions—opportunities that allow them time, flexibility, and the experience of competence. Learners in these phases of interest development already notice and wonder and have begun to ask and seek answers to curiosity questions. They want to think more about the content itself (e.g., the connections between what they presently understand and new ways of thinking about this information—e.g., that the worms need soil to live). In the transition from less to more developed phases of interest, learners appear to become promoters and developers of their own interest. They also need the kind of support and challenge that comes from working with or alongside others, through feedback, identifying new goals, and so forth.

As the research reviewed and the snapshots of this chapter suggest, it appears that interest for particular content can and will develop and deepen when a learner's phase of interest, feelings of self-efficacy, and ability to self-regulate are both acknowledged and used to provide ongoing support to recognize, take advantage of, and seek opportunities to think and reengage.

Clarification about how the others and tasks of the learning environment might provide such support is still needed, however. This is a charge for both research and practice.

ACKNOWLEDGMENTS

I would like to acknowledge Louise Chawla's recanting of Jane Goodall's talk and Ming Cai's description of Optimus Prime, the thoughtful insight and support of Jessica Bachrach, Melissa Emmerson, Suzanne Hidi, Alex List, David Preiss, and Robert Sternberg, and the help of Nadine Kolowrat with editing. Work on this chapter was supported by the Swarthmore College Faculty Research Fund.

NOTES

1. Interest is always linked to some class of objects, activities, or events; this may be a domain (e.g., science) or a more focused topic (e.g., dissections in biology). In the present discussion, *interest* should be understood as possibly referring to both. However, the existence, process, and impact of interest as a psychological state and predisposition to engage is the focus of the present discussion, not the particular content (e.g., science vs. soccer).

Differences among domains of potential interest do exist (e.g., some are more hierarchical and require mastery of skills in order to progress, whereas others allow for a number of different types of connections), and these may affect the nature of engagement as well as the possibilities for learning (e.g., mathematics, chess, and music are more hierarchical, whereas English literature, biology, and psychology are less so). Research addressing interest for particular content in which distinctions between types of interest based on domain demands is reported in Breen and Lindsay (2002); Johnson, Alexander, Spencer, Leibham, and Neitzel (2004); Krapp and Fink (1992); Rheinberg and Schiefele (1997).

2. Hidi, Renninger, and Krapp (2004) describe four of these differences. They discuss them in relation to other motivational variables. Discussion of interest as potentially unreflective represents an extension of that discussion.

3. The potential that learner motivation may differ based on disciplinary content is recognized (e.g., Urdan & Schoenfelder, 2006); this line of theorizing is only a recent development in the motivation research. It provides implicit support for the thesis of the present chapter: that learners differ in their interest and also in their feelings of self-efficacy and ability to self-regulate.

4. A related concept is the self-concept of ability; *self-concept of ability* refers to discipline-based rather than specific task-based feelings of ability (see Bong & Skaalvik, 2003). For studies of interest and self-concept of ability, the reader is pointed to studies by Denissen, Zarrett, and Eccles (2007) and Nieswandt (2007).

5. The snapshots draw on different research projects undertaken by Renninger and colleagues. The interpretation of the relation among interest, self-efficacy, and self-regulation in each of the snapshots is aided by case analyses of learners in other studies as well, including those reported in Lipstein and Renninger, 2007b; Renninger, 2000; Renninger, 2009; Renninger et al., 2008; Renninger, Ewen, and Lasher, 2002; Renninger and Hidi, 2002; and Renninger, Sansone, and Smith, 2004.

REFERENCES

Alexander, P. A. (2004). A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp.273–298). Mahwah, NJ: Erlbaum.

Alexander, P. A., Kulikowich, J. M., & Schulze. (1994). The influence of knowledge, domain knowledge, and interest on the comprehension of scientific exposition. *Learning and Individual Differences, 6,* 379–397.

Armstrong, P. I., Allison, P., & Rounds, J. (2008). Development and initial validation of brief public domain RIASEC marker scales. *Journal of Vocational Behavior, 73,* 287–299.

Ashby, F. G., Iseni, A. M., & Turken, A. U. (1998). Profiling the differences in students' knowledge of positive affect and its influence on cognition. *Psychological Review, 106,* 529–550.

Azevedo, F. S. (2006). Personal excursions: Investigating the dynamics of student engagement. *International Journal of Computers for Mathematical Learning, 11,* 57–98.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

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

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development, 49,* 193–224.

Barron, K. E., & Harackiewicz, J. M. (2000). Achievement goals and optimal motivation: A multiple goals approach. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 229–254). New York: Academic Press.

Bergin, D. A. (1999). Influences on classroom interest. *Educational Psychologist, 34,* 87–92.

Boekaerts, M. (2006). Self-regulation and effort investment. In K. A. Renninger & I. E. Sigel (Vol. Eds.), *Child psychology in practice* (Vol. 4, pp. 345–377). In W. Damon & R. M. Lerner, (Gen. Eds.), *Handbook of Child Psychology* (6th ed.). New York: Wiley.

Bong, M., & Skaalvik, E. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15,* 1–40.

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *Journal of Social Psychology, 130,* 353–363.

Bouffard-Bouchard, T., Parent, S., & Larivee, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *International Journal of Behavioral Development, 14,* 153–164.

Breen, R., & Lindsay, R. (2002). Different disciplines require different motivations for student success. *Research in Higher Education, 43*, 693–725.

Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology, 50,* 492–501.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88,* 715–730.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S. (1993). *Talented teenagers: The roots of success and failure.* New York: Cambridge University Press.

Deci, E. L. (1992). The relation of interest to the motivation of behavior: A Self-Determination Theory perspective. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 43–70). Hillsdale, NJ: Erlbaum.

Denissen, J. H., Zarrett, N. R., & Eccles, J. S. (2007). "I like to do it, I'm able, and I know I am": Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78,* 430–447.

Dewey, J. (1913/1975). *Interest and effort in education.* Boston: Riverside Press.

Dweck, C. S. & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review, 95,* 256–273.

Eccles, J. P., & Midgley, C. (1989). State/environment fit: Developmentally appropriate classrooms for early adolescents. In R. Ames & C. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 139–181). New York: Guilford Press.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64,* 830–847.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, pp. 1017–1095). New York: Wiley.

Elliot, A. J., & Harackiewicz, J. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70,* 461–475.

Flowerday, T., & Schraw, G. (2003). Effect of choice on cognitive and affective engagement. *Journal of Educational Research, 96,* 207–215.

Flum, H., & Kaplan, A. (2006). Exploratory orientation as an educational goal. *Educational Psychologist, 41*(2), 99–110.

Ford, M. (1992). *Motivating humans: Goals, emotions, and personal agency beliefs.* Newbury Park, CA: Sage.

Gisbert, K. (1998). Individual interest in mathematics and female gender identity: Biographical case studies. In L. Hoffmann, A. Krapp, K. A. Renninger, & J. Baumert (Eds.), *Interest and learning: Proceedings of the Seeon conference on interest and gender* (pp. 387–401). Kiel, Germany: IPN.

Goodall, J. (March, 2008). *Roots and shoots.* Presentation at the Garfield Park Conservatory, Chicago.

Hannover, B., & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction, 14,* 51–67.

Harackiewicz, J., Barron, K. E., Carter, S. M., Lehto, A. T., & Elliot, A. J. (1997) Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology, 73,* 1284–1295.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S.M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement: Predicting continued interest and performance over time. *Journal of Educational Psychology, 92,* 316–330.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94,* 562–575.

Harackiewicz, J. M., & Elliot, A. J. (1993). Achievement goals and intrinsic motivation. *Journal of Personality and Social Psychology, 65,* 904–915.

Harackiewicz, J., & Sansone, C. (1991). Goals and intrinsic motivation: You can get there from here. *Advances in Motivation and Achievement, 7,* 21–49.

Harter, S. (1985). Competence as a dimension of self-evaluation: Toward a comprehensive model of self-worth. In R. Leahy (Ed.), *The development of the self* (pp. 55–121). New York: Academic Press.

Harter, S. (1986). Processes underlying the construction, maintenance and enhancement of the self-concept in children. In J. Suls & A. C. Greenwald (Eds.), *Psychological perspectives on the self* (pp. 137–181). Hillsdale, NJ: Erlbaum.

Harter, S. (1999). *The construction of the self.* New York: Guilford Press.

Harter, S. (2003). The development of self-representation during childhood and adolescence. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 610–642). New York: Guilford.

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 311–342). New York: Academic Press.

Hidi, S. (2006). Interest: A motivational variable with a difference. *Educational Research Review, 1,* 69–82.

Hidi, S., & Ainley, M. (2008). Interest and self-regulation: Relationships between two variables that influence learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp.77–109). Hillsdale, NJ: Erlbaum.

Hidi, S., & Boscolo, P. (2006). Motivation and writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 144–197). New York: Guilford.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70,* 151–179.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41,* 111–127.

Hidi, S., Renninger, K. A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 89–115). Mahwah, NJ: Erlbaum.

Hoffmann, L. (2002). Promoting girls' learning and achievement in physics classes for beginners. *Learning and Instruction, 12,* 447–465.

Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology, 6,* 35–45.

Husman, J., & Lens, W. (1999). The role of the future in student motivation. *Educational Psychologist, 43,* 113–125.

Johnson, K. E., Alexander, J. M., Spencer, S., Leibham, M. E., & Neitzel, C. (2004). Factors associated with the early emergence of intense interests within conceptual domains. *Cognitive Development, 19*, 325–343.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review, 19,* 429–422.

Kessels, U. (2005). Fitting into the stereotype: How gender stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education, 20,* 309–323.

Kilpatrick, J., Swafford. J., & Findell, B. (Eds.). (2002). *Adding it up.* Washington, DC: National Academy Press.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Kitsantas, A., & Zimmerman, B. J. (2002). Comparing self-regulatory processes among novice, non-expert, and expert volleyball players: A microananalytic study. *Journal of Applied Sport Psychology, 14,* 91–105.

Krapp, A. (2002). Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. *Learning and Instruction, 12,* 383–409.

Krapp, A., & Fink, B. (1992). The development and function of interests during the critical transition from home to preschool. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 71–98). Hillsdale, NJ: Erlbaum.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology, 38,* 424–430.

Lepper, M. R., & Cordova, D. L. (1992). A desire to be taught: Instructional consequences of intrinsic motivation. *Motivation and Emotion, 16,* 187–208.

Lepper, M. R. & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 257–307). New York: Academic Press.

Linnenbrink, E. A., & Pintrich, P. R. (2000). Multiple pathways to learning and achievement: The role of goal orientation in fostering adaptive motivation, affect, and cognition. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 196–227). New York: Academic Press.

Lipstein, R., & Renninger, K. A. (2007a). "Putting things into words": 12–15-year-old students' interest for writing. In P. Boscolo & S. Hidi (Eds.), *Motivation and writing: Research and school practice* (pp. 113–140). New York: Kluwer Academic/Plenum Press.

Lipstein, R., & Renninger, K. A. (2007b). Interest for writing: How teachers can make a difference. *English Journal, 96,* 79–85.

Maehr, M., & Midgley, C. (1991). Enhancing student motivation: A school wide approach. *Educational Psychologist, 26,* 399–427.

Marsh, H. W. (1991). The failure of high ability high schools to deliver academic benefits: The importance of academic self-concept and educational aspirations. *American Educational Research Journal, 28,* 445–480.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology, 85,* 424–436.

Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching, 44*(7), 908–937.

Nolen, S. B. (2006). The role of literate communities in the development of children's interest in writing. In P. Boscolo & S. Hidi (Eds.), *Motivation and writing: Research and school practice* (pp. 241–256). New York: Kluwer Academic/Plenum Press.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology, 21,* 325–344.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86,* 193–203.

Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching, 46*(2), 147–165.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotion.* New York: Oxford University Press.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13,* 423–451.

Perry, T. (2003). Up from the parched earth: Toward a theory of African-American achievement. In T. Perry, C. Steele, & A. Hillard (Eds.), *Young, gifted, and black: Promoting high achievement among African American students* (pp. 1–108). Boston: Beacon Press.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16,* 385–407.

Pintrich, P. R. & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 249–284). New York: Academic Press.

Prenzel, M. (1992). The selective persistence of interest. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 71–98). Hillsdale, NJ: Erlbaum.

Pressick-Kilborn, K., & Walker, R. (2002). The social construction of interest in a learning community. In D. M. McInerney & S. Van Etten (Eds.), *Sociocultural influences on motivation and learning: An historical perspective* (pp. 153–182). Greenwich, CT: Information Age.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337–386.

Renninger, K. A. (1990). Children's play interests, representation, and activity. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 147–172). Norwood, NJ: Ablex.

Renninger, K. A. (1992). Individual interest and development: Implications for theory and practice. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development.* 361–395 Hillsdale, NJ: Erlbaum.

Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 375–407). New York: Academic Press.

Renninger, K. A. (2009). Interest and identity development in instruction: An inductive model. *Educational Psychologist, 44*(2), 105–118.

Renninger, K. A., Bachrach, J. E., & Posey, S. K. E. (2008). Learner interest and achievement motivation. In M. L. Maehr, S. Karabenick, & T. Urdan (Eds.), *Social psychological perspectives, volume 15: Advances in motivation and achievement* (pp. 461–491). Bingley, UK: Emerald Group.

Renninger, K. A., Bachrach, J. E., Riley, K. R., Niwagaba, A. A. P., & Tibbets, C. Y. (April, 2009). Exploration and interest: Inner-city young adolescents engagement in an out-of-school science inquiry workshop. Paper presented as part of the symposium, *Exploration as a motivational construct: Forging a research agenda.* American Educational Research Association, San Diego, CA.

Renninger, K. A., Ewen, E., & Lasher, A. K. (2002). Individual interest as context in expository text and mathematical word problems. *Learning and Instruction, 12,* 467–491.

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 173–195). New York: Academic Press.

Renninger, K. A. & Lipstein, R. (2006). Come si sviluppa l'interesse per la scrittura; Cosa volgliono gli studenti e di cosa hannobisogno? [Developing interest for writing: What do students want and what do students need?] *Età Evolutiva, 84,* 65–83.

Renninger, K. A., Sansone, C., & Smith, J. L. (2004). Love of learning. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification* (pp. 161–179). New York: Oxford University Press.

Renninger, K. A., & Shumar, W. (2004). The centrality of culture and community to participant learning at and with The Math Forum. In S. Barab, R. Kling, & J. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 181–209). New York: Cambridge University Press.

Renninger, K. A., & Shumar, W. (2002). Community building with and for teachers: *The Math Forum* as a resource for teacher professional development. In K. A. Renninger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (pp. 60–95). New York: Cambridge University Press.

Rheinberg, F., & Schiefele, U. (1997). Motivation and knowledge acquisition: Searching for mediating processes. *Advances in Motivation and Achievement, 10*, 251–301.

Roeser, R. W., Peck, S. C., & Nasir, N. S. (2006). Self and identity processes in school motivation, learning, and achievement. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 391–424). Mahwah, NJ: Erlbaum.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 66–78.

Sansone, C. (1986). A question of competence: The effects of competence and task feedback on intrinsic interest. *Journal of Personality, 74,* 1697–1720.

Sansone, C. & Smith, J. L. (2000). Interest and self-regulation: The relation between having to and wanting to. In C. Sansone, & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 341–372). New York: Academic Press.

Sansone, C. Weir, C., Harpster, L., & Morgan, C. (1992). Once a boring task always a boring task? Interest as a self-regulatory mechanism. *Journal of Personality and Social Psychology, 63,* 379–390.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review, 13,* 23–52.

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32,* 195–208.

Schunk, D. H., & Zimmerman, B. J. (2008). *Motivation and self-regulated learning: Theory, research, and applications.* New York: Taylor & Francis.

Schwartz, D. L., & Bransford, J. D. (1998). *Cognition and Instruction, 16*(4), 475–522.

Silvia, P. J. (2006). *Exploring the psychology of interest.* New York: Oxford University Press.

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44,* 331–349.

Vygotsky, L. (1978). *Mind in society.* Boston: Harvard University Press.

White, R. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66,* 297–333.

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In *The development of achievement motivation* (pp. 91–120). New York: Academic Press.

Wigfield, A., & Harold, R. D. (1992). Teacher beliefs and children's achievement self-perceptions: A developmental perspective. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 95–121). Hillsdale, NJ: Erlbaum.

Yamauchi, L. A., Wyatt, T. R., & Carroll, J. H. (2005). Enacting the five standards for effective pedagogy in a culturally relevant high school program. In A. E. Maynard & M. I. Martini (Eds.), *Learning in cultural contexts: Family, peers, and schools* (pp. 227–245). New York: Kluwer.

Zahorik, J. A. (1996). Elementary and secondary teachers' reports of how they make learning interesting. *The Elementary School Journal, 96*(5), 551–564.

Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal, 31,* 845–862.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainments: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29,* 663–676.

Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process to outcome goals. *Journal of Educational Psychology, 8,* 29–36.

Zimmerman B. J. & Kitsantas, A. (2006). A writer's discipline: The development of self-regulatory skills. In S. Hidi & P. Boscolo (Eds.), *Motivation and interest in writing* (pp. 51–72). Amsterdam: Elsevier.

Zimmerman, B. J., & Schunk, D. H. (2004) Self-regulating intellectual processes and outcomes: A social cognitive perspective. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 323–349). Mahwah, NJ: Erlbaum.

PART III

Refining Mind

This page intentionally left blank

6

Finding Young Paul Robesons: Exploring the Question of Creative Polymathy

JAMES C. KAUFMAN, RONALD A. BEGHETTO, AND JOHN BAER

Paul Robeson lived many lives. Born in 1927, he was a superstar college athlete across many sports, a Bo Jackson or Jim Thorpe, who could excel at basketball and track while also serving as class valedictorian (Boyle & Buni, 2005). Advancing to law school, Robeson supported himself by being a professional football player and an actor. Leaving the world of sports and law, he became a celebrated singer and actor. He played Joe in both the stage and film versions of *Show Boat,* singing the definitive "Ol' Man River" (perhaps think of the way that a Frank Sinatra or Barbra Streisand might claim a song for their own). His portrayal of Othello in London and on Broadway was the first major interpretation of the role by an African American (and considered quite controversial at the time). His performance was acclaimed in the same way that Laurence Olivier would be praised years later. Indeed, *Othello,* based on his now famous interpretation, is today typically performed by African American actors. Even if Robeson had stopped adding to his accomplishments at that point, his eminence in these four very different arenas (scholarship, athletics, acting, and music) would warrant this introductory paragraph.

Yet Robeson continued to be a creative force in new areas. He traveled to the Soviet Union and Europe as part of his singing and acting career, and he spoke to international audiences about the prejudice shown to African Americans. He became a civil rights activist, both for

African Americans (he founded a national movement to combat lynching) and other groups, such as Asian Americans and blue collar workers (Foner, 2004). In part because of anticommunist government policies and in part because of blatant racism, the FBI targeted Robeson (much as they would later target Martin Luther King, Jr.). Having been accused of un-American activities, he lost his passport and was subjected to frequent persecution; some prominent African Americans later told of being asked to denounce Robeson in exchange for not being targeted themselves (Boyd, 2004). Ill health and political fallout kept Robeson from being as active during the last decades of his life; he died in 1978. He won several international peace awards during his lifetime as well as countless posthumous honors for his activism, music, and athletic achievements.

People who excel creatively in so many different areas stand out because of their rarity. Imagine someone alive today who possessed the athletic talent of Tiger Woods, the acting ability of Kate Winslet, and the voice of Josh Groban—*and* was a strong advocate of social justice and peace along the lines of Jimmy Carter, Nelson Mandela, or Vaclav Havel. Waiting for the handful of women and men who can join the ranks of Leonardo Da Vinci, Benjamin Franklin, Clare Booth Luce, Bertrand Russell, and Linus Pauling, however, can be an exercise in disappointment. Such people are rare. Does the rarity of highly accomplished creative polymaths (i.e., those who are creative in more than one domain) mean that only certain people have multicreative potential? We argue in this chapter that the scarcity of large numbers of multidomain geniuses does not necessarily mean that creative polymathy is impossible or even highly unlikely, and we make some tentative suggestions of ways in which we might nurture the multicreative abilities of students.

We develop our argument by first considering multicreative potential in light of the "Four C" model (Beghetto & Kaufman, 2007; Kaufman & Beghetto, 2009). Next we examine arguments that contest whether creativity is (a) a *general* construct (i.e., a model arguing that the same processes that allow someone to be creative in one domain also enable that person to be creative in others, so that the skills that lead to creative performance are the same or very similar across all domains) or (b) a very *domain-specific* construct (i.e., a model claiming that the skills underlying creative performance vary from domain to domain, so that the skills that help one be creative in one domain would be of little use in other, unrelated domains). We offer an alternate possibility to these two extreme models by drawing on the Amusement Park Theoretical (APT)

Model (Baer & Kaufman, 2005; Kaufman & Baer, 2004, 2006), a hierarchical model that draws on both generality and specificity in a way that would predict some (but not unlimited) polymathy. We then present a marionette analogy in which someone might be creative in one domain through one set of cognitive, social, or environmental processes and then might also be creative in a different domain through a potentially overlapping yet substantially different set of processes. Finally, we discuss how each person's creative marionette can vary depending on his or her level of creative development.

THE FOUR-C MODEL OF CREATIVITY

One obstacle to truly understanding the nature of multicreative ability is the very conception of what it means to be creative. Typically, there are two approaches to studying creativity: larger or "Big-C" creativity and smaller or "little-c" creativity (Csikszentmihalyi, 1996). *Big-C* focuses on eminent creativity. The goals of this approach are often to learn about creative genius and discuss which creative works may last forever (e.g., Simonton, 1994). Big-C creators are extremely prodigious in their achievements, have been recognized as revolutionary by gatekeepers of a domain (Csikszentmihalyi, 1996), and typically have devoted many—generally 10 or more—years of intense study developing the domain expertise necessary for making a revolutionary contribution (Ericsson, Roring, & Nandagopal, 2007; Simon, 1981; Simonton, 1997). Creative greatness may be studied by analyzing the lives of well-known creators, interviewing renowned individuals, or studying people who excel at high levels on creativity measures. Big-C creativity has traditionally been the focus of much research and theorizing in creativity studies. However, an equally long-standing tradition has been to caution against an overly narrow focus on eminent creative productivity. Stein (1953), for instance, argued that when scholars focus on genius-only creativity, it places too great an emphasis on the more objective, externally referenced forms of creativity and obscures the equally important—yet more subjective or internally referenced—experiences of creativity.

Fortunately, creativity researchers, in recent decades, have expanded conceptions to include more everyday or little-c experiences of creativity (Richards, 1990). This little-c work explores the creativity inherent in the daily activities and experiences in which the average person may participate (e.g., coming up with a new bedtime story for a son or daughter that meaningfully incorporates the events of the past day; combining leftovers into an unexpected and tasty fusion of flavors; or finding a creative solution to the problem of how to deal with an unpleasant person at work).

Although the distinction between Big-C and little-c creativity is useful, it has important limitations—as is true with most dichotomies. In particular, the two categories do not sufficiently account for subtle yet meaningful distinctions in creative experience and expression. Consider, for instance, a poet who has published several chapbooks of poetry but may never attain the status of an eminent poet (whose work is anthologized and taught in poetry classes). This poet clearly is not a Big-C poet, but to categorize her as a little-c poet would diminish her professional success (as she would be lumped in with the occasional holiday poet who writes poetry for family and friends and may not be interested in or able to sell his poetry to a publishing house). The professional poet, considered in light of the Big-C/little-c dichotomy, is misplaced or obscured. The same can be said for the much more subjective creativity of a student learning how to write poetry. Although the student may have new and personally meaningful insights about the use of imagery in her poetry, such insights may not be sufficiently novel to be considered creative even at the little-c level. Given these limitations with the traditional Big-C/little-c dichotomy, Kaufman and Beghetto (2009) propose a *Four-C Model of Creativity* to include two additional categories: "mini-c" creativity (Beghetto & Kaufman, 2007) and "Pro-c" creativity (Kaufman & Beghetto, 2009).

Mini-c creativity includes the "personal" (Runco, 1996) and "developmental" (Cohen, 1989) aspects of creativity. Mini-c creativity pertains to subjective self-discoveries—the novel and personally meaningful insights and interpretations inherent in the learning process. Mini-c may blossom when one solves an algebraic equation, cracks a joke, tells a story, or builds a Lego tower. Although mini-c creativity may not meet the traditional standards used for Big-C or even little-c, at the very least it should be considered a sign of creative potential. As Vygotsky (2004, original work published in 1967) has argued, internal creative acts can still be considered creative, even when they only take the form of "some mental or emotional construct that lives within the person who created it and is known only to him" (p. 7).

A fourth-grade student writing his first Haiku poem is a good example of mini-c creativity. Although his poem may not bring anything "new" to the arena of Haiku writing and may not even adhere completely to the conventions of the style (for example, his use of syllables might not be perfect), there is nonetheless a core of internal creative expression taking place. For instance, he is discovering new ways of using the constraints of Haiku writing to represent personally meaningful images and emotions associated with the falling of leaves. Such mini-c insights and interpretations, while valuable in their own right, can also serve as building blocks for the expression of further creative insights (e.g., the creation of a prizewinning Haiku poem at the county fair).

Pro-c creativity refers to professional-level creators who have not yet attained truly eminent status (Kaufman & Beghetto, 2009). Someone who writes folk songs in his spare time falls into the little-c category, and Bob Dylan is a likely example of Big-C. The Pro-c singer–songwriter, in contrast, would be someone who has played in paying venues and may even have released a few CDs of music but is not necessarily close to leaving a permanent mark on the field. Professionals in the various academic disciplines (geologists, computer science professors, English literature critics) would also be considered examples of pro-c creators.

APPLYING THE FOUR-C MODEL TO THE CREATIVE POLYMATH

The Four-C model helps broaden conceptions of creativity to include everything from the more subjective, mini-c creative insights and potential to the more objective and clear-cut examples of Big-C creative eminence. How might this model be helpful in nurturing the multicreative potential of youngsters? One way in which it is helpful is that it allows researchers to consider the likelihood of expressing multicreative potential across the various levels of creative magnitude (Beghetto & Kaufman, 2009) and thereby helps educators to consider, more realistically, how they might nurture the multicreative potential of their students.

If, for instance, the goal is to nurture polymathy at the Big-C level, educators are faced with a nearly impossible challenge. Multidomain Big-C creators (such as Robeson, Da Vinci, and Franklin) are extremely rare. This is not surprising given what it takes to achieve creative eminence in just one domain—recognition by critics, historians, or other relevant gatekeepers of a domain—not to mention the many years of intensive study and even some degree of chance (Simonton, 2004). This confluence of factors makes it very difficult to predict whether a particular youngster will be capable of a Big-C contribution in one domain, let along across multiple domains. Even being a child prodigy—while increasing the likelihood of later-life creative accomplishment—provides no guarantees (Howe, 1999).

Of course, the fact that achievement at eminent levels of creative accomplishment is extremely rare does not mean that educators should give up on attempts to nurture their students' creative potential. Rather, as the Four-C model helps to show, it simply means that educators might need to set their sights on identifying and nurturing the much more likely multicreative accomplishments at the smaller-c levels. A good place for educators to start is by increasing their own (and their students') awareness of individuals who have arrived at multicreativity across the various levels of creative magnitude.

Root-Bernstein (1989), for instance, has documented more than 400 cases of famous scientists who were also highly skilled—at the pro-c or little-c level—in the arts. Consider, for example, Galileo, who is best known for his Big-C scientific work but was also a skilled musician and painter. Even from simply an anecdotal perspective, Pro-c creative polymaths are relatively common in quite distinct domains. Consider also John Glenn (astronaut and politician), Arnold Schwarzenegger (actor and politician), Brian May (guitarist and astrophysicist), and the late Paul Newman (actor and entrepreneur). The floodgates open up when we allow a mix of Pro-c and little-c. Many (if not most) accomplished academics that we know (especially those who study creativity) have little-c avocations in such areas as poetry, stand-up comedy, art, theater, photography, music, and business.

Finding someone at the Pro-c level who is also creative at an everyday level in a different domain is quite easy. When the category becomes a pairing of little-c (or better) creative accomplishments, polymaths become common. At the mini-c level, one is probably more likely to find creative polymathy than not; part of the human developmental process is learning and trying many different types of activities and tasks. Mini-c polymathy can occur any time someone combines new and personally meaningful insights and interpretations across different disciplines or domains.

Considering creative polymathy across the Four Cs of creativity can help educators recognize how they might better support the multicreative potential of all students in their classrooms. For instance, teachers might encourage their students to explore creative connections within and across academic disciplines—such as asking students to incorporate their own unique and meaningful insights about the design principles they are learning in art class into the development of poster presentations for the school science fair. Giving youngsters various opportunities to explore the process of creativity—within and across disciplines—is in accordance with what Root-Bernstein (2003) has reported as being a common childhood experience of accomplished polymaths.

In order to offer students such opportunities, educators must attempt to strike a balance between providing breadth and depth in such experiences (Plucker & Beghetto, 2004). This poses a nontrivial challenge, which becomes more or less pronounced as a result of many factors including the pedagogical expertise of teachers; the curricular supports and constraints placed on teachers; and the interests, experiences, and varying levels of competence that students bring into the classroom. Providing a breadth of disciplinary opportunities may, for instance, be more challenging but no less important for students at a specialized school like Julliard; whereas providing depth of experiences at an ordinary neighborhood school becomes challenging because advanced or enrichment classes are not readily available. Still, to the extent that teachers are able to make some progress in striking this balance, they put themselves in a better position to support the potentials that might otherwise remain hidden.

DOMAIN SPECIFICITY VERSUS DOMAIN GENERALITY: WHAT DOES IT MEAN FOR POLYMATHY?

Polymathy at the little-c level, we have argued, is quite common. However, it is here that we run into evidence that is seemingly contradictory, because the research on creativity across multiple domains (at the little-c level) suggests that creativity may be very domain-specific. Does this doom our argument for little-c polymathy? To many people, domain specificity *does* seem to rule out polymathy, because if the skills that underlie creativity are completely different in different domains, then a person's skills leading to creativity in one area would be of no value in other domains. How then, people ask, could one person be creative in multiple domains? We think this is a mistaken argument, but to answer it we must first try to understand what research does in fact say about domain specificity.

Here is how one creativity researcher has argued that the question of whether creativity is domain-specific or domain-general might be adjudicated:

> Domain generality would be supported by high intercorrelations among different creative behaviors and a common set of psychological descriptors for those behaviors, while domain specificity would be supported by relatively low correlations among different behaviors, and a diverging set of psychological descriptors of those behaviors. (Ivcevic, 2007, p. 272)

One possible way to answer this question—a way that has been used by several researchers—is to ask subjects to produce creative products in a variety of domains and then have expert judges in those different domains rate the creativity of the results in their respective domains. The researchers can then look to see if the subjects who produced more creative works in one domain were more likely to produce creative works in other domains. To do this they simply look for correlations across the ratings in different domains. Baer, for example, has tested students ranging from second-graders to college students. He had these students produce creative work through writing poetry, writing short stories, telling stories out loud, creating mathematical equations, creating mathematical word problems, and making collages. He consistently found low and usually nonsignificant correlations between ratings of creative performance in these different areas (Baer, 1991, 1993, 1994a, 1994b, 1996). In other words, a student who wrote a creative poem was *not* more likely to also tell a creative story, make a creative collage, or write a creative mathematical equation (a creative algebraic equation might use numbers in a playful or unusual way). Several other studies (e.g., Han, 2003; Runco, 1989) have found similar results. And if you remove variations due to IQ, the small correlations get even smaller.

Even in studies that have claimed to show higher degrees of domain generality, closer examination has shown very little evidence of domain generality. For example, Conti, Coon, and Amabile (1996) analyzed data from studies in which subjects had both written stories and engaged in art activities. The intercorrelations among the creativity ratings of the stories were high, confirming the prediction that "creativity measures within the same domain are substantially intercorrelated" (p. 387). Intercorrelations of creativity ratings among the art tasks, which were more unlike one another than were the story-writing tasks, were positive but somewhat lower. But within-domain correlations tell us nothing about the domain specificity–generality question. Cross-domain correlations, on the other hand, speak directly to the generality–specificity question. None of the 13 cross-domain correlations that Conti and colleagues reported—the crucial tests for domain generality—was statistically significant, and the mean of these 13 correlations was just .109, accounting for just a little more than 1% of the variance.

Other kinds of evidence argue for domain generality, however, and not everyone is convinced by the evidence favoring domain specificity (see, e.g., Plucker, 1998, 2005; Plucker & Beghetto, 2004). Research that looks at actual creative products tends to yield results like those just cited favoring domain specificity, but other methods can and do support domain generality. As Plucker (1998) has argued, "the conclusions of researchers using [performance measures of creativity] are almost always that creativity is predominantly task or content specific . . . Conversely, researchers utilizing traditional psychometric methods usually conclude that creativity is predominantly content general" (p. 181).

The kinds of psychometric evidence to which Plucker refers are primarily self-report scales. (Divergent-thinking tests, such as the Torrance Tests of Creative Thinking, cannot help here because they *assume* domain generality and therefore offer no way to evaluate the possibility of domain specificity.) Self-report scales tend to support at least a modest degree of domain generality. For example, Hocevar (1976) found "low to moderate" (p. 869) correlations among self-report indexes of creativity in various domains among college students. In a study in which several thousand subjects self-reported their own creativity in 56 domains, Kaufman, Cole, and Baer (2009) found both an overarching general factor and seven more specific areas of creative performance.

The issue of domain specificity–generality is, then, an open question in creativity research. Because domain specificity would seem to argue against the likelihood of finding polymaths, however, we need to explain why even if the domain specificity theorists are right, this would still *not* rule out the possibility (or even likelihood) of creative polymathy. With that concern out of the way, we can then present our model, which includes features of both the domain-general and domain-specific approaches, and explain how polymathy fits into this wider conception of creativity.

WHY DOMAIN SPECIFICITY WOULD *NOT* MEAN THAT NO ONE IS MULTITALENTED

Even if one accepts the most extreme version of the claim that creativity is domain-specific, that does not mean that polymaths are either impossible or unexpected. Domain specificity does not claim that *no one* has a multitude of creative abilities (just as it does not claim that everyone is creative in one domain or another). It simply says that (a) the skills that underlie creativity vary by domain and (b) the presence or absence of any particular skill or set of skills (or the degree to which these skills exist, because they are not dichotomous, either–or kinds of abilities) in any one individual involves a degree of randomness (e.g., chance encounters, experiences, or opportunities that support or undermine the development creativity relevant skills in a domain).

Consequently, a small fraction of people will have impoverished or severely underdeveloped creativity-relevant skills in just about any domain; many people will have developed modest amounts of skills in several domains; some will have developed a great deal of skill in one or more domains; and a few will have developed great quantities of such skills in many domains. Here is an analogy: If there were 1,000 each of red, blue, green, and orange marbles that were randomly distributed among 100 people, a few people might end up with no marbles of any color and a few others might end up with several dozen marbles of every color. Most people would get some mix, which might be a modest number of marbles of all colors or lots of marbles of some colors and few of other colors. That is how randomness works. Of course, because the fact that the development of creativity-relevant skills involves some degree of randomness does not mean that underdeveloped creativity in a particular domain can never be developed. Rather, understanding the role that chance plays in the development of creativity-relevant skills helps explain the distribution of people who have (or have not) developed the skills requisite for creative expression within or across domains.

Paul Robeson was creative in many domains. That may have been (under a domain-general interpretation) because he had a great deal of creativity, which he could apply to whatever field interested him at the time. But under a domain-specific interpretation, Robeson's creativity in many domains could just as easily be explained by claiming that he happened to have many creativity-relevant abilities in many domains. The talents that led to his success as a singer need not be the same as (or even overlap with) the skills that helped him to be a great athlete, and neither set of skills might have had anything to do with his success as a student, an actor, or a civil rights activist. There may have been overlap, but there need not have been. Someone can be talented in math and also be a good tennis player, and yet these may be entirely distinct domains that are based on completely different underlying abilities. Being creative in two seemingly unrelated areas does not show that creativity is domain-general any more than the existence of a mathematically talented tennis expert would prove that tennis and math are rooted in the same set of skills. Domain specificity, even in its most extreme form, does not argue that people can be creative in only a single domain. It simply argues that because the abilities that make creativity possible in different domains are different, creative performance in one domain does not predict creativity in other areas. If domain-based talents are randomly distributed, then one should find a few people who have a great deal of creativity-relevant skills in many domains, some people who have talents (of varying degrees) in several domains, and some who have little talent in any domain. This is what a normal distribution of unrelated skills would predict. So the presence of a few polymathic "Renaissance men" would not contradict domain specificity. In fact, it is precisely what domain specificity predicts.

Just as the presence of a few extraordinary individuals like Paul Robeson leads some people to think (mistakenly) that the existence of such people is evidence against domain specificity, the rarity of such polymaths sometimes leads people (also mistakenly) to think that the fact that polymaths are rare argues against domain generality. Why, they ask, if creativity is domain-general, did Albert Einstein not also write wonderful poems and Emily Dickinson not produce great paintings? The answer here—as we have already suggested—is what psychologists call the "10-year rule" (Hayes, 1989). The 10-year rule claims that it takes at least 10 years of preparation before "even the most noteworthy and 'talented' individuals" (Weisberg, 1999, p. 230) can develop the kinds of knowledge and skill needed to produce genius-level work in any domain. As Gruber and Davis (1988) wrote, "Perhaps the single most reliable finding in our studies is that creative work takes a long time" (p. 264). Because it takes at least 10 years of preparation in a given domain to reach the highest levels of creative accomplishment in that domain, it should come as little surprise that few people manage to reach those levels in more than one (or perhaps two or three at most) fields in a single lifetime. Kaufman and Kaufman (2007) studied 215 contemporary novelists. They found that the writers took an average of 10.6 years between their first publication and their best publication. Just as past research indicates that it takes 10 years to be published as an expert after starting to put the pen to the paper, there may also be evidence that another 10 years must pass before a truly elite work is produced.

Singer/songwriter Joni Mitchell is also a very talented painter and has said "I have always thought of myself as a painter derailed by circumstance" (Kelly, 2000). She once visited artist Georgia O'Keefe, who told that "I would have liked to have been a painter and a musician, but you can't do both." Mitchell replied, "O yes, you can!" (Weller, 2008, p. 427). The point that O'Keefe was making—that it is *very* hard to devote oneself and one's time and energies to more than one field—is the same point that the 10-year rule makes. Mitchell's response is a reminder that polymathy at the highest levels of accomplishment in at least two very different domains is possible but certainly not easy.

So under either domain specificity or domain generality, one would expect to see at least a few but not a large number of multitalented individuals who do paradigm-shifting work in many fields. But neither model rules out the possibility of there being *many* people who might be Big-C creative in one domain and also Pro-c creative in several others, or simply Pro-c creative in many fields (under the assumption that the 10-year rule limits Big-C creativity far more than Pro-c creativity). And domain specificity certainly does *not* lead one to expect a scarcity of polymaths at the little-c or mini-c levels. (Domain generality would lead us to expect polymathy to be even more widespread than would domain specificity, but neither theory makes specific predictions.) In the next section we present a model that can provide a framework for the many kinds and degrees of creativity that we see in the world—a model that we believe can help us understand both single talents *and* polymathy. It can also help us to find promising approaches for identifying talented students.

THE AMUSEMENT PARK THEORETICAL MODEL OF CREATIVITY

A theory that can lead us (in conjunction with the Four C Model) to new insights about creative domains is the Amusement Park Theoretical (APT) Model of Creativity. Because the details of the model are presented elsewhere (Baer & Kaufman, 2005; Kaufman & Baer, 2004, 2006), the key features of this theory are briefly summarized to demonstrate how it might make possible the kind of testing in the realm of creativity that we have begun to see in intelligence testing.[1] The APT model is based (somewhat whimsically, perhaps, as some of our reviewers have noted) on a large amusement park. In an amusement park there are initial requirements (e.g., a ticket) that apply to all areas of the park. Similarly, there are initial requirements that, to varying degrees, are necessary to creative performance in all domains (e.g., intelligence, motivation).

Amusement parks also have *general thematic areas* (e.g., at Disney World one might select from EPCOT, the Magic Kingdom, the Animal Kingdom, and Disney–MGM Studios), just as there are several different general areas in which someone could be creative (e.g., the arts, the sciences). Once in one type of park, there are sections (e.g., Fantasyland, Tomorrowland), just as there are *domains* of creativity within larger *general thematic areas* (e.g., physics and biology, domains in the *general thematic area* of science). These domains, in turn, can be subdivided into *microdomains* (e.g., in Fantasyland one might visit Cinderella's Castle or It's a Small World; in the domain of psychology, one might specialize in cognitive psychology or social psychology).

As an example, if one were interested in assessing the creative abilities of a subject in terms of poetry writing, one might start by assessing such *initial requirements* as a certain minimal level of intelligence as well as skill in the *general thematic area* of language. Next, one might assess skills in certain *domains* especially relevant to writing poetry (e.g., metaphor-generating ability; see Baer, 1996, for other examples of such domain-specific skills related to creativity in poetry and how such skills can be trained). Finally, if one was interested only in the ability to write haiku (and not other kinds of poetry), one might evaluate skills in specific micro-domains related only to that kind of poetry.

If, on the other hand, one were interested in a student's creative potential in the area of physical science, the hierarchy of skills that one would evaluate would be quite different. The *initial requirements* might be similar (e.g., a certain level of intelligence), but skills from very different *general thematic areas* would be of interest (e.g., verbal skills would be less important, and the ability to understand and to generate unusual mathematical ideas would be of much greater interest). The differences would become even greater as one moved down the hierarchy to *domains* and *microdomains.*

Motivation could also be assessed at different levels of such a hierarchy. For example, a student might have strong intrinsic motivation at the level of the *general thematic area* of science, and this would indicate a tendency toward creative productivity in the sciences in general. Another student might have extremely high intrinsic motivation only in the *domain* of marine science, however, which predicts a greater likelihood of creativity in that domain but not in other sciences. Or a student's interest at a given point in time might be even more narrowly focused on a *microdomain* (for example, a student might have great interest in the reproductive success of certain kinds of mollusks in different environments but show little interest in other areas of marine science).

This suggests that while polymathy could occur (and *should* at times occur) with respect to almost any set of abilities, polymathy would be most likely to occur with respect to different abilities in the same general thematic area. In combination with the Four-C model, we would expect the greatest number of polymaths to exhibit abilities in the same general thematic area but at different levels (and especially not all at the highest or Big-C level). We would also expect a somewhat smaller number of polymaths who exhibited high levels of creativity in very different general thematic areas (e.g., the artistic–verbal, artistic–visual, entrepreneurial, interpersonal, math/science, performance, and problem-solving areas).

Here is a more concrete example: The APT and Four-C Models would predict that more people would exhibit a pattern of Big-C creativity as novelists, Pro-c creativity as poets, and little-c creativity in journalism and drama than would either (a) exhibit Big-C creativity in all four areas or (b) would exhibit Big-C creativity as novelists, Pro-c creativity as physicists, and little-c creativity as sculptors and dancers—but either theory would not completely rule out either of these latter possibilities. They are simply much less likely to occur than a mixture of levels of creativity in domains in the same general thematic area.

STRANGE BEDFELLOWS

Even within the fairly structured confines of the APT Model, there will also always be strange bedfellows that make sense only upon closer examination. A person might decide to pick his favorite amusement park based only on how good the popcorn is at the food court. Someone else may go only to cheap amusement parks (Big Alan's Generic Roadside Attraction). In a similar way, microdomains or domains may be selected for reasons that are less than obvious. Maybe Phil does not have a lot of money. He therefore pursues creative domains that do not call for a big outlay of funds, such as poetry, stand-up comedy, and geology (he finds interesting rocks and looks for patterns in their shapes).

We believe that many interesting creative polymaths can be found hidden within these strange connections. Certainly, in examining people who are creative in two microdomains, one will typically find the creativity in microdomains that are included in the same domain. Examples might be a journalist who is equally creative in writing editorials and movie reviews or a psychologist who can creatively plan experiments in both cognitive and social psychology. However, it is the people who are creative in dissimilar areas—like Robeson, whose accomplishment included athletics, acting and singing, and activism—who represent the pinnacle of creative polymathy.

It is the same phenomenon behind the strange bedfellows in the APT Model that brings us to the distinction between being creative across multiple domains versus creative polymathy. Most studies that have examined the domain-specific versus domain-general question have picked their domains; so, for example, Baer (1993) has given students such tasks as story writing, storytelling, poetry writing, the creation of mathematical equations and word problems, and collage making. He found consistently low and usually nonsignificant correlations between creative ability in these different areas. In other words, a student who wrote a creative poem was *not* more likely to also tell a creative story or write a creative mathematical equation. The key point behind most of these studies is to argue that creativity is not a general construct that will manifest itself across all areas.

The domain-specific point of view does not insist, as explained above, that people be allowed to be creative in only one domain. Instead, it suggests that the underlying components of creativity are probably different from one domain to another. Let us say that Jacob is creative at formulating algebraic proofs and at tap dancing. Someone arguing the general party line of pure creativity would say that the elements enabling Jacob to be creative in both areas are the same. The domain-specific approach would consider such double proficiency the equivalent of being able to both bench press 300 pounds and recite pi to two hundred places—both neat things to be able to do but feats that are based on quite different abilities!

Imagine, however, a puppeteer manipulating two marionettes. She is making the first marionette play the piano and the second marionette juggle. Both marionettes are engaged in activities and they certainly share a commonality—the same person is pulling the strings for both. But the strings themselves are completely different. We believe that this marionette analogy can hold for creative polymaths. Whatever components enabled Paul Robeson to be a creative singer (such as being able to hold a tune, enjoy music, and reflect emotion in his voice) may have been different than those components that enabled him to be a creative political activist (such as being able to empathize with others, organize others, and have the courage to stand for his convictions). And yet he (Paul Robeson) was himself the *puppeteer* who was responsible for his creative *marionettes* in those different domains.

Like the APT Model, the puppeteer/marionette analogy allows us to consider multiple levels of creative talent or skill. There may be some general abilities (the APT Model's initial requirements) that influence creativity in many areas, just as the skill and dexterity of the puppeteer makes it possible for him to manipulate many different kinds of marionettes. But that general puppeteering skill gets one nowhere without specific marionettes, each of which has its own strings to pull and its own possible range of performances. In the same way, an initial requirement like general intelligence may be important for creativity in many areas, but general intelligence alone is not enough. One also needs more specific skills and motivations in particular general thematic areas and specific domains if one is to evidence creativity. And this is true whether one is thinking about genius-level Big-C creativity, high-level Pro-c creativity, or more everyday little-c or even mini-c creativity (although of course the necessary degree of domain-specific talent is far greater at the higher levels of creativity).

IDENTIFYING AND NURTURING TALENT

How can we use these models to help us identify students who may be polymaths (or who may have specific creative talents with or without polymathy)? First, it is important to remember that we should not expect young students to find Big-C or Pro-c creativity. Their little-c and mini-c creativity may someday become Big-C creativity, but not all Big-C creators display the extraordinary precocity of a Mozart, who at age 5 could read, write, and play music proficiently. There is no one-to-one correspondence between precocity and later outstanding achievement. As Ellen Winner (1996), an authority on highly gifted children, wrote, "Only a very few of the gifted become eminent adult creators. We cannot assume a link between early giftedness, no matter how extreme, and adult eminence" (p. 11).

Focusing on our students' little-c and mini-c creativity is not to suggest that we should overlook any young Mozarts in our classes (something almost impossible to do) but rather to emphasize that it is the less extreme (and less obvious) little-c and mini-c talents and interests that we should be working hard to identify. The APT Model does not claim that there are no general creativity-relevant skills, just as a good identification program for gifted/talented students should not ignore IQ test scores; but it reminds us that there are many far less general abilities and motivations that are just as important. Young students who show an extraordinary joy in playing with words; children who find delight in making sketches of scenes from the books they love (and who may have what Winner (1996) called a "rage to draw" (p. 87) in her description of one talented young artist); students who devour books about stars, architecture, or some other area of special interest: these are all potentially creatively talented children. Any special skill or interest may be evidence of creativity in a given domain or microdomain. Our search for giftedness and talent should not be limited to those who have high IQs (although those with high IQs should certainly be included as well), and our efforts to nurture talent should, to the extent possible, harmonize with the specific talents that our students exhibit.

Can we expect to find students who exhibit creative polymathy? Yes and no. Like Paul Robeson, our students may exhibit their talents at different times rather than all at once. Robeson's talents as a civil rights activist were not apparent early in his athletic career. But many of his achievements in diverse fields did overlap, and so they sometimes will in our students. To the extent possible, we are wise to follow our students' creative muses, which may have their own calendars and schedules. If a student's interests include both music and science, we would do well to nurture both; but if only one talent or special interest is apparent at a given time, then that is the area on which we should focus.

Once they have been identified, what can we do to nurture budding talents and special interests? There are at least two kinds of things that the psychology of creativity tells us are helpful: (1) show interest but get out of the way and (2) help students develop domain-relevant skills and knowledge. Here's what we mean.

Show Interest but Get Out of the Way

Expressing interest in children's ideas, projects, activities, performances, and passions is helpful. Talking with them (and especially listening to them talk) about their interests and providing resources they might need (materials, books, tools, instruments, contacts, etc.) are also important. But then we often need to get out of their way. Intrinsic motivation is a wonderful thing. Unfortunately, when extrinsic motivation is added, the net result may be temporarily higher total motivation, but in the long term it often leads to lower intrinsic motivation. Bribing children (or adults) to do things they already like to do seems to turn fun into work, and extrinsic motivation more generally (e.g., rewards, anticipated evaluation) has this unfortunate long-term effect while at the same time often decreasing creativity (Amabile, 1983, 1996; Baer, 1997b, 1998; Hennessey & Amabile, 1988; Lepper & Greene, 1975, 1978). This effect is not a simple one, and there is dispute in the creativity research community about its generality. There is evidence that rewards that are more clearly targeted or tied to specific behaviors can either have no negative effect or may sometimes even enhance creativity (Eisenberger, Pierce, & Cameron, 1999; Eisenberger & Shanock, 2003). But the evidence that extrinsic motivators can decrease both intrinsic motivation and creativity is far too strong to ignore, and the last thing a teacher wants to do, even inadvertently, is to take away a student's joy or passion for learning.

Perhaps the most common kind of reward/bribe that teachers use with talented students is extra credit. The teacher's motivation is often a positive one. For example, after noticing that a student has a real passion for astronomy, the teacher may offer extra credit for making a model of the solar system; but the likely effect is not to increase the student's interest but rather (somewhat counterintuitively) to reduce it. We all like getting rewards, and what could be better than getting extra credit for doing something one likes? The problem is that such well-intentioned bribes tend to have very negative long-term consequences. What was originally something the student did because of genuine interest has become something she does to get points from her teacher, who has partially taken over direction of her pursuit. If the student who loved astronomy had wanted to make a model of the solar system, she would have done so without the extra credit. Instead she was pursuing her interest in astronomy in her own way. Motivated by the lure of extra credit, however, she has given over some control of her previously self-directed study and allowed the teacher to redirect her efforts to do something that is perhaps not as interesting but which—unlike her previous self-directed effort—is rewarded. Her teacher has, quite unintentionally, converted fun into work.

This is not to say that all extrinsic motivation is bad. Students need to learn skills; in order to learn skills, they need feedback on their performance. Learning to anticipate evaluation is an essential skill. Teachers must give their students—including their most creative and talented ones—the kinds of feedback they need to develop their skills. Such feedback should be informative and improvement-focused—letting students know what they have done well, helping students develop their understanding of domain conventions and constraints, and pointing out how they might continue to improve and develop their ideas, insights, and interpretations in light of the particular domain and task constraints (Beghetto, 2007). In providing such feedback, teachers should also try to minimize social comparison, stress the informative aspects of their evaluative feedback, and acknowledge intellectual risk taking inherent in students' creative expression (Beghetto, 2005).

It is worth bearing in mind that it is generally far easier to kill intrinsic motivation than to instill it, and teachers should be careful not to use rewards or evaluations as tools to motivate students in areas where they already have significant levels of intrinsic motivation. It is often wiser just to get out of the way of such highly motivated students than to try to promote or encourage interests that need no promotion (Baer, 1997a).

Help Students Develop Domain-Relevant Skills and Knowledge

It takes many years to learn everything that one might need to know to become an expert (and possibly a Big-C creator) in any field, as noted above. Students therefore need opportunities to learn about the domains that are of particular interest to them, and they cannot wait until those things happen to appear (or fail to appear) in the regular curriculum. If students have special interests and/or talents, we can nurture those interests by providing opportunities for students to learn more, far more than the curriculum generally expects, in their areas of special interest.

This assistance might take the form of arranging for them to participate in special programs (e.g., summer science camps) or simply providing books and other resources. It might involve special training (e.g., voice lessons) or connecting students with others (both peers with similar interests and abilities and adults who might become mentors) who share their passions. Opportunities to learn about the lives (both professional and personal) of Big-C creators in a student's field of interest can also be helpful.

For students who have many special interests and talents—little-c polymaths—teachers might also provide help with time management. A student with many special interests can never have enough time for all she wants to do, not to mention those pesky things that the world seems to think she *has* to do (e.g., those other subjects—the ones that might not, at least not yet, excite her).

Polymaths like Paul Robeson are rare, and we are unlikely to encounter many (or any) in our teaching careers. But Big-C creators often have many Pro-c or little-c talents and interests (Root-Bernstein &

Root-Bernstein, 2004), and many students exhibit vary degrees of little-c and mini-c creativity in diverse domains. One can never know which talents will become the most important to a polymath at any point in his or her life. Joni Mitchell thought music would remain a hobby, playing second fiddle to her serious work as a painter (DiMartino, 1998), and Paul Robeson surely could not have known which of his many talents would become most important to him later in life. So it is with our students. What we can offer them is the chance to develop any and all of their creative talents.

NOTE

1. The APT Model is not the only one to address domain specificity and generality; see also Plucker and Beghetto's (2004) Hybrid model.

REFERENCES

Amabile, T. M. (1983). *The social psychology of creativity.* New York: Springer-Verlag.
Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity.* Boulder, CO: Westview.
Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal, 4,* 23–39.
Baer, J. (1993). *Divergent thinking and creativity: A task-specific approach*. Hillsdale, NJ: Erlbaum.
Baer, J. (1994a). Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Research Journal, 7,* 35–46.
Baer, J. (1994b). Generality of creativity across performance domains: A replication. *Perceptual and Motor Skills, 79,* 1217–1218.
Baer, J. (1996). The effects of task-specific divergent-thinking training. *Journal of Creative Behavior, 30,* 183–187.
Baer, J. (1997a). *Creative teachers, Creative students.* Boston: Allyn & Bacon.
Baer, J. (1997b). Gender differences in the effects of anticipated evaluation on creativity. *Creativity Research Journal, 10,* 25–31.
Baer, J. (1998). The case for domain specificity in creativity. *Creativity Research Journal, 11,* 173–177.
Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The Amusement Park Theoretical (APT) Model of creativity. *Roeper Review, 27*, 158–163.
Beghetto, R. A. (2005). Does assessment kill student creativity? *The Educational Forum, 69,* 254–263.
Beghetto, R. A. (2007). Ideational code-switching: Walking the talk about supporting student creativity in the classroom. *Roeper Review, 29,* 265–270.
Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for mini-c creativity. *Psychology of Aesthetics, Creativity, and the Arts, 1,* 73–79.

Beghetto, R. A., & Kaufman, J. C. (2009). Do we all have multicreative potential? *ZDM: The International Journal of Mathematics Education, 41,* 39–44.

Boyd, H. (2004, October 28). McCarthyism revisited and dissected. *New York Amsterdam News, 95,* 10, 44.

Boyle, S. T., & Buni, A. (2005). *Paul Robeson: The years of promise and achievement.* Amherst: University of Massachusetts Press.

Cohen, L. M. (1989). A continuum of adaptive creative behaviors. *Creativity Research Journal, 2,* 169–183.

Conti, R., Coon, H., & Amabile, T. M. (1996). Evidence to support the componential model of creativity: Secondary analyses of three studies. *Creativity Research Journal, 9,* 385–389.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention.* New York: HarperCollins.

Eisenberger, R., Pierce, W. D., & Cameron, J. (1999). Effects of reward on intrinsic motivation: Negative, neutral, and positive. *Psychological Bulletin, 125,* 677–691.

Eisenberger, R., & Shanock, L. (2003). Rewards, intrinsic motivation, and creativity: A case study of conceptual and methodological isolation. *Creativity Research Journal, 15,* 121–130.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies, 18,* 3–56.

Foner, H. (2004). Forward—Keynote Address from the 1998 Long Island University Paul Robeson Conference. In J. Dorinson & W. Pencak (Eds.), *Paul Robeson: Essays on his life and legacy* (pp. 1–5). Jefferson, NC: McFarland.

Gruber, H. E., & Davis, S. N. (1988). Inching our way up Mt. Olympus: The evolving-systems approach to creative thinking. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 243–270). New York: Cambridge University Press.

Han, K. S. (2003). Domain-specificity of creativity in young children: How quantitative and qualitative data support it. *Journal of Creative Behavior, 37,* 117–142.

Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Roning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 202–219). New York: Plenum.

Hennessey, B. A., & Amabile, T. M. (1988). Conditions of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 11–38). Cambridge: Cambridge University Press.

Hocevar, D. (1976). Dimensionality of creativity. *Psychological Reports, 39,* 869–870.

Howe, M. J. A. (1999). Prodigies and creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 431–446). Cambridge: Cambridge University Press.

Ivcevic, Z. (2007). Artistic and everyday creativity: An act-frequency approach. *Journal of Creative Behavior, 41,* 271–29.

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, J. C., & Baer, J. (2006). Intelligent testing with Torrance. *Creativity Research Journal, 18,* 99–102.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The Four C Model of Creativity. *Review of General Psychology, 13,* 1–12.

Kaufman, J. C., Cole, J. C., & Baer, J. (2009). The construct of creativity: A structural model for self-reported creativity ratings. *Journal of Creative Behavior, 43,* 119–134.

Kaufman, S. B., & Kaufman, J. C. (2007). Ten years to expertise, many more to greatness: An investigation of modern writers. *Journal of Creative Behavior, 41,* 114–124.
Kelly, D. (2000, August 6). I sing my sorrow and I paint my joy. *Toronto Globe and Mail,* p. R1.
Lepper, M., & Greene, D. (1975). Turning work into play: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. *Journal of Personality and Social Psychology, 31,* 479–486.
Lepper, M., & Greene, D. (1978). *The hidden costs of reward.* Hillsdale, NJ: Erlbaum.
Plucker, J. A. (1998). Beware of simple conclusions: The case for the content generality of creativity. *Creativity Research Journal, 11,* 179–182.
Plucker, J. A. (2005). The (relatively) generalist view of creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 307–312). Hillsdale, NJ: Erlbaum.
Plucker, J. A., & Beghetto, R. A. (2004). Why creativity is domain general, why it looks domain specific, and why the distinction doesn't matter. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 153–168). Washington, DC: American Psychological Association.
Richards, R. (1990). Everyday creativity, eminent creativity, and health: "Afterview" for CRJ issues on creativity and health. *Creativity Research Journal, 3,* 300–326.
Root-Bernstein, R. (1989). *Discovering*. Cambridge, MA: Harvard University Press.
Root-Bernstein, R. (2003). The art of innovation: Polymaths and universality of the creative process. In L. V. Shavinina (Ed.), *The international handbook on innovation* (pp. 267–280). Oxford: Elsevier.
Root-Bernstein, R., & Root-Bernstein, M. (2004). Artistic scientists and scientific artists: The link between polymathy and creativity. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 127–152). Washington, DC: American Psychological Association.
Runco, M. A. (1989). The creativity of children's art. *Child Study Journal, 19*, 177–190.
Runco, M. (1996). Personal Creativity: Definition and developmental issues. *New Directions for Child Development*, 72, 3–30
Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.
Simonton, D. K. (1994). *Greatness: Who makes history and why.* New York: Guilford Press.
Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104,* 66–89.
Simonton, D. K. (2004). Creativity as a constrained stochastic process. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 83–101). Washington, DC: American Psychological Association.
Stein, M. I. (1953). Creativity and culture. *The Journal of Psychology, 36,* 311–322.
Vygotsky, L. S. (2004). Imagination and creativity in childhood. (M. E. Sharpe, Inc., Trans.) *Journal of Russian and East European Psychology, 42,* 7–97. (Original work published 1967)
Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 226–250). New York: Cambridge University Press.
Weller, S. (2008). *Girls like us: Carole King, Joni Mitchell, Carly Simon—and the journey of a generation.* New York: Atria.
Winner, E. (1996). *Gifted children: Myths and realities.* New York: Basic Books.

7 How (Well-Structured) Talk Builds the Mind

LAUREN B. RESNICK, SARAH MICHAELS, AND M. C. O'CONNOR

Language is, historically and individually, the foundation of being human. And *talk*—direct exchange between humans who can attribute intentionality and understanding to each other—is the foundational act of language. Without talk, we cannot achieve full humanity or social community. Without talk, minds can neither grow nor become disciplined. Without disciplined talk, scientific, mathematical, and humanistic knowledge remains static and unused. Without disciplined talk, society withers and raw conflict among individuals, groups, and nations becomes an increasingly dangerous way of life.

How does talk actually work—for social decision making, and for learning complex academic disciplines? And how might we expand both the individual and societal capacity for using language productively? These are the questions guiding this chapter, which considers several examples of productive talk, showing first how the rules that govern human conversation often mask fundamentally rational processes that seem to be learnable by everyone. Turning then to academic disciplines, we show what productive classroom talk in the disciplines looks like. A surprising finding we explore is that lack of fundamental knowledge about issues

This work was supported in part by the Pittsburgh Science of Learning Center, which is funded by the National Science Foundation, award number SBE-0354420.

under discussion—rather than an inability to reason well—often blocks successful learning through talk. Having shown that certain forms of talk can nevertheless enhance academic learning, we then consider evidence that learning in this way can produce quite broad "transfer"—that is, effects beyond the discipline directly taught. At the end we draw some broad conclusions about what would be needed to make effective talk-based learning the norm in schools and in society more generally.

DELIBERATIVE DEMOCRACY

Globalization, multiculturalism, and diversity—whether ethnic, racial, or socioeconomic—now require new approaches to social life and decision making. In an increasingly connected but diverse world, deliberation and discussion must be employed not just to communicate what people already know or believe, but also to build knowledge and craft negotiated solutions to complex political, medical, and environmental problems. An emerging body of work in philosophy, political science, psychology, and linguistics addresses these issues on both theoretical and practical grounds. A common point of reference, reaching across academic specializations, is Jurgen Habermas's (1990) notion of "deliberative democracy" and the "public sphere as an idealized discursive space where debate and dialogue are free and uncoerced." The idea of deliberative democracy has been taken up by a wide range of political and legal theorists who see deliberative democracy as a productive response to both liberalism (emphasizing the rights and freedoms of the individual) and communitarianism (emphasizing group solidarity and identity).

Dialogue and discussion have long been linked to theories of democratic education. Before Habermas, from Socrates to Dewey, educative dialogue was promoted as a forum for learners to develop understanding by listening, reflecting, proposing, and incorporating alternative views. In fact, Dewey proposed a definition of democracy that placed reasoned discussion at its very heart. He spoke of democracy as a "mode of social inquiry" emphasizing discussion, consultation, persuasion, and debate in the service of just decision making (Dewey, 1966).

Philosophers and psychologists working in the informal logic tradition have also focused their attention on the processes of informal argument (Blair & Johnson, 1987; van Eemeren & Grootendorst, 1988; Voss, Perkins, & Segal, 1991; Walton & Krabbe, 1995). Although philosophers have clarified the rational norms underlying different kinds of discussion, it has been the specific goal of cognitive researchers to capture the implicit rules and constraints that guide the social construction of reasoning in particular contexts (Anderson, Chinn, Chang, Waggoner, & Yi, 1997).

REASONED DISCUSSION AND THE NORMS OF CONVERSATION

What does "reasoned discussion" look like in ordinary human conversation? With colleagues from philosophy and linguistics, one of us (Resnick) set out almost 20 years ago to study this question (Resnick, Salmon, Zeitz, Wathen, & Holowchak, 1993). The path was arduous and the results surprising in multiple ways. We began by recording conversations and creating verbatim transcripts of university students discussing controversial public policy issues such as the expansion of nuclear power facilities and the permissibility of prayer in public schools. We brought together students in groups of three, having previously assessed their views on the issues so we could ensure diversity of starting opinion in the discussion. We asked them to try to arrive at an agreement in the course of about 20 minutes of discussion. There was no appointed leader, and discussion proceeded as a conversation rather than a formal debate.

We assumed at the beginning that we could construct and apply a relatively simple coding system based on the theory of informal reasoning that had been persuasively developed by Stephen Toulmin (1958). Toulmin, attempting to move beyond the constraints of formal logic as a criterion of good reasoning, had proposed a structure of *claims*, *warrants*, and *backings* as the structures of reasoning. We quickly learned that conversational norms made it essentially impossible to directly apply a structured coding based on Toulmin's analysis. People did not speak in complete grammatical sentences, much less reasoned arguments that included explicit claims, warrants, and backings. In fact, our first attempts at analysis almost led us to abandon the effort to find rationality in the discussion. Fortunately, with the help of a multidisciplinary (psychology, philosophy, linguistics, education) team of faculty and students, we persevered.

We ultimately developed a graphic coding system that took apart the utterances of individuals and traced themes across multiple contributions and over time (Resnick et al., 1993). Our analytic scheme took into account sociolinguistic tenets of conversation (Grice, 1968). Conversational analysis was then showing that when people converse, they try both to support the positions they espouse (in Toulmin's terms, they provide warrants and backings) and to respond sensitively to others in the conversation. Furthermore, the rules of conversation implicitly constrain individuals from overtly saying everything they mean. Conversationalists rely on their partners to "fill in the blanks" to make their contributions make sense. Indeed, saying everything would be conversationally improper; it would take too long and fail to respect the intelligent participation of the other (Resnick, Levine, & Teasley, 1991).

Exhibit 7.1 contains a transcript of the first 2 minutes of one group's discussion on nuclear power. We began our analysis by building a simple display that allowed us to examine the ways in which the participants in the conversation distribute the thematic content of their shared problem-solving efforts.

Exhibit 7.1

TWO MINUTES OF DISCUSSION ON NUCLEAR POWER

A1:	Who's gonna go first?
A,B,C:	(laughter)
B2:	Well you're "A" so why don't you go first
A3:	Oh boy, well I don't know . . So what do you//
C4:	Well, uh, is is nuclear, I'm against it . . Is nuclear power really cleaner than fossil fuels? I don't think so
A5:	You don't think. I think//
B6:	In terms of atmospheric pollution I think that . . the waste from nuclear power, I think it's . . much less than fossil fuels. . but the waste that is there of course is quite dangerous//
C7:	It's gonna be here for thousands of years, you can't do anything with it. I mean, right now we do not have the technology as//
A8:	Acid rain lasts a long time too you know
C9:	That's true but if you reduce the emissions of fossil fuels which you can do with, uh, certain technology that we do have right now, um, such as scrubbers and such, you can reduce the acid rain, with the nuclear power you can't do any, I mean nuclear waste you can't do anything with it except//

--(1 min.)--

TWO MINUTES OF DISCUSSION ON NUCLEAR POWER *(Continued)*

B10:	bury it
A11:	m-hm
C12:	bury it and then you're not even sure if its ecologically um . . that the place you bury it is ecologically sound.
B13:	I, I think if if enough money is spent it can probably be put in a reasonably safe area
A14:	reasonable for what? (laugh)
C15:	well//
B16:	well//
C17:	Well we just heard that out in New Mexico//
B18:	If something, something is either half or three quarters of a mile underground in a geologically solid area, no earthquake or anything like that is going to disturb it and get it into the water table
C19:	You see but the only problem is the burial is gonna have to be for a hundred of thousand years//
B20:	m-hm
C21:	And you cannot guarantee that anything is gonna be geologically safe for a hundred thousand years. . at least I don't think I can do that, or maybe even, maybe you can but it doesn't seem possible to me
B22:	m-hm
A23:	Well I suppose it's a problem . . I really, I really don't know//
C24:	right
A25:	that much about geology//
B26:	m-hm
A27:	Enough what uh how long things stay stable for . . you know.

--(2 min.)--

Note: Two periods indicate a pause of between 1 and 2 seconds. Double backslashes indicate places where a speaker is interrupted.

Figure 7.1 shows this display for part of the transcribed piece of conversation. The analysis in Figure 7.1 is deliberately superficial, in the sense that it stays very close to the surface structure of the transcription of the conversation, with little concern for interpreting or discovering features of reasoning. Utterances with no thematic content that seemed to play no guiding role in the conversation were eliminated. Otherwise, the participants' language was maintained, more or less verbatim, but with some compression in order to fit the graphic constraints.

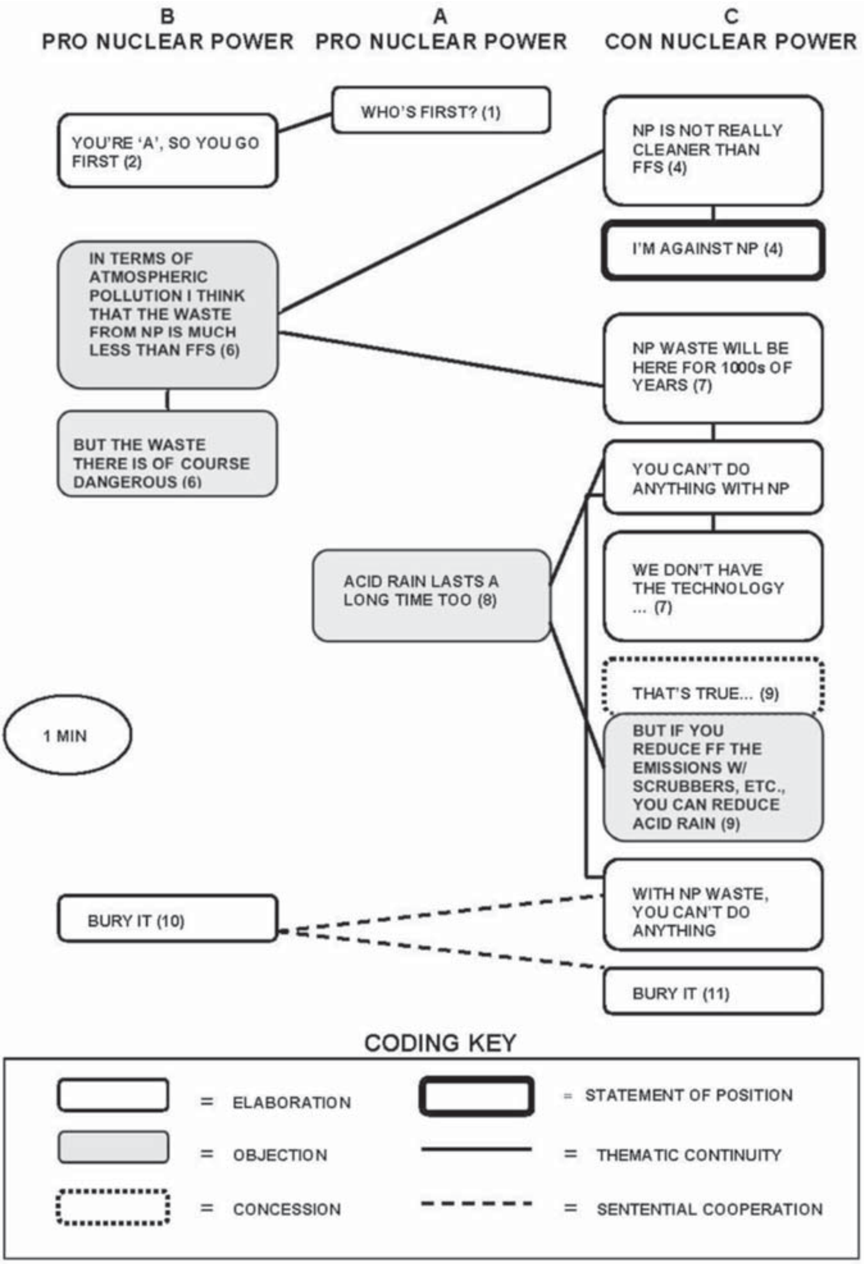


Figure 7.1 Thematic content of shared problem-solving efforts.

To generate Figure 7.1, participants' speech was broken up into idea units corresponding roughly to noncompound sentences, one box per unit. Time is represented vertically. The sequence of idea units is distributed among three columns, one for each participant. Lines connect the boxes that appear to have thematic connections. These judgments were straightforward; verbal and nonverbal cues (from the videotapes we made of the discussions) made it a low-inference task to decide when people were referring to their own or their coparticipants' previous utterances. Only three judgments about content were made: clear boxes were judged to be elaborations or repetitions of the content of the earlier boxes to which they are connected, shaded boxes express an objection to or disagreement with the content of an earlier box, and dotted boxes indicate a concession that one participant made to the (expressed or implied) opinion of another. These, too, were low-inference judgments. A concession box linked together with a disagreement box indicates a "yes . . . but" structure, a form that occurs frequently in all our data. Sentential cooperation (one speaker completing the idea of another) is shown by dotted lines connecting idea units. Even in this very brief protocol and its graphic analysis, we can see two of the most important features of shared argument development. First, arguments are distributed across participants, jointly constructed. People "finish each other's arguments," so to speak. Second, argument construction is distributed over time.

However, a third and very important feature of the conversational argument became apparent only when we developed a further analysis of the argument structure. Figure 7.2 suggests the complexity. Shapes are used to mark the argumentation function of each utterance (conclusion, implied conclusion, premise, etc.), and sets of utterances (shapes) can be grouped into arguments and attacks on arguments.

Figure 7.3 shows a graphic display of just a short exchange between two of the participants in the transcript. This form of analysis revealed that certain ideas that are *not overtly stated* can play a crucial role in the conversation. These are *unstated premises.* For example, in a discussion of sources of power, discussants might imply that "the only two sources of power are nuclear and coal." Such an implied premise might be challenged by a discussion partner (that is what participant B was doing in the Figure 7.3 segment). If the challenge is accepted by the original speaker (participant C in the example), we have confirmation that the premise was indeed part of the argument. On examining many analyses such as the one in Figure 7.3, we learned that our conversation

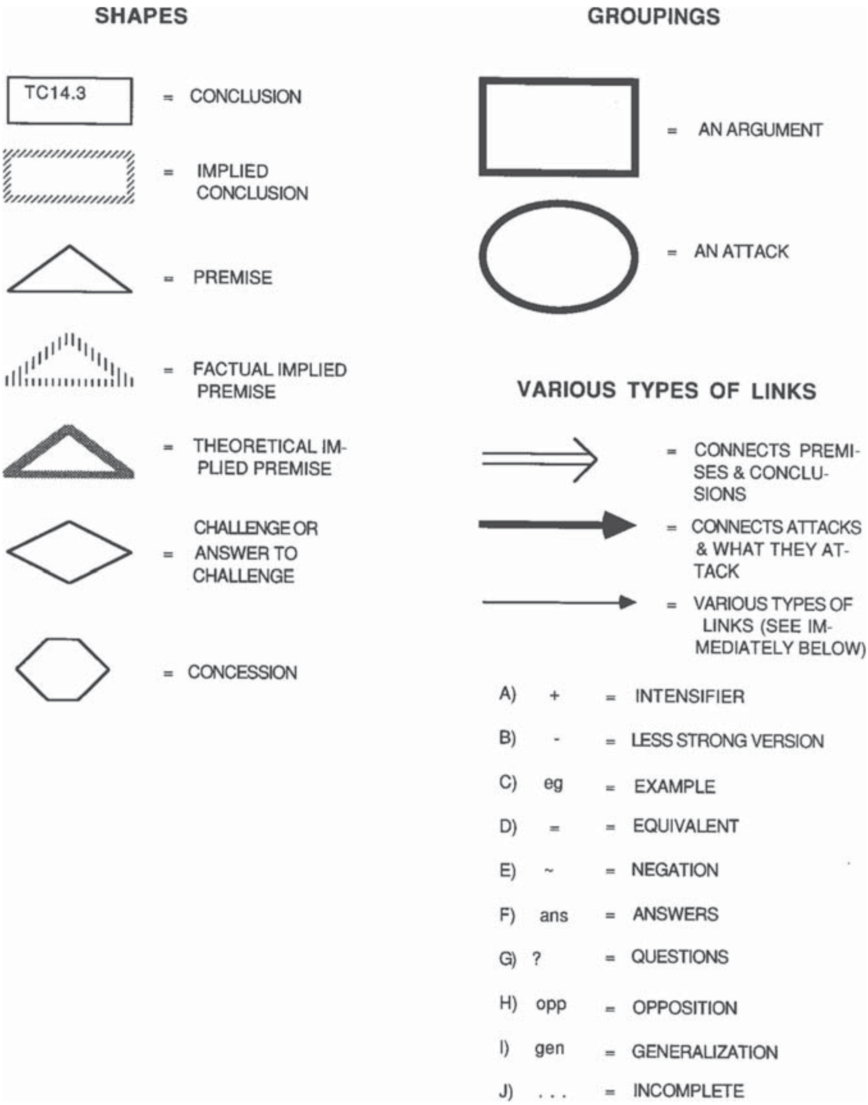


Figure 7.2 Shapes used to represent types of argumentation moves.

groups built complex argument and attack structures, often challenging premises rather than directly attacking conclusions. These exchanges moved the conversation beyond confrontation of opinions or simple acceptance of disagreement. New arguments were actually built in the course of conversation.

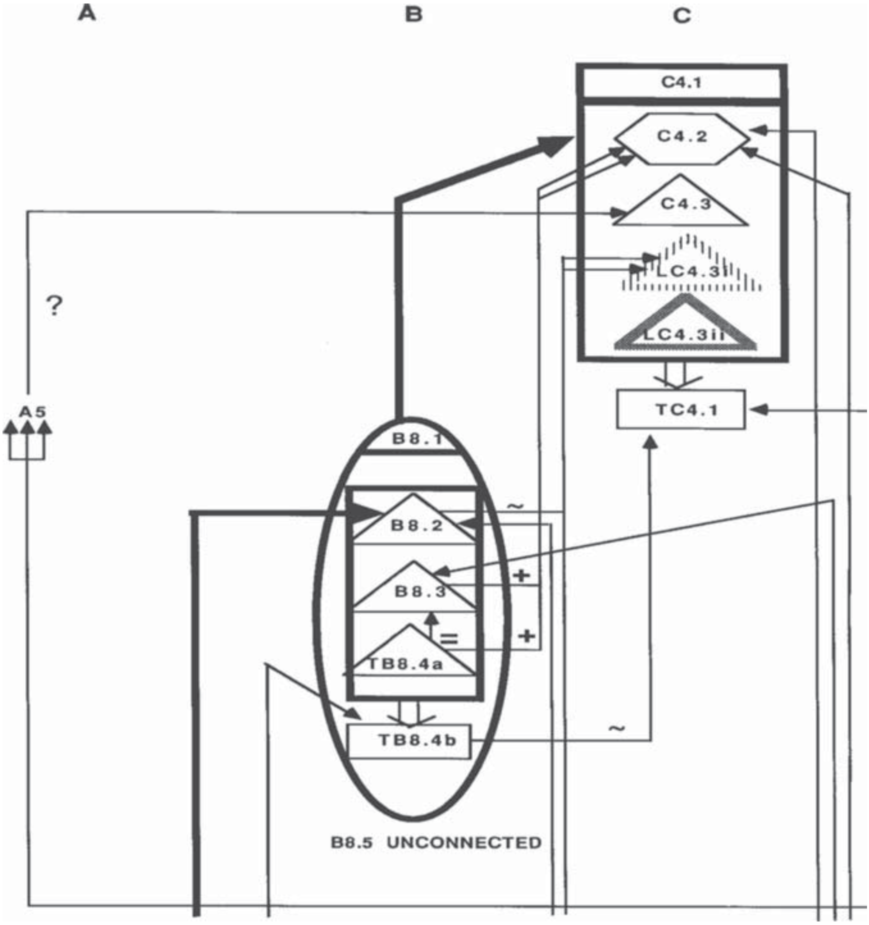


Figure 7.3 Graphic display using argument move shapes reveals unstated premises as in LC4.3 in the top right of Column C.

LEARNING DELIBERATIVE DISCOURSE

How are the kinds of reasoning-in-conversation skills displayed in the previous example learned? Had students been explicitly taught them? Were these exceptionally smart students? We cannot answer this question with certainty at present, but many who have been studying the functioning of language in social contexts and more general aspects of the human mind are coming to believe—or at least entertain the hypothesis—that the basic capacity to use language to reason is universally human. It might be somehow "in the genetic code," so to speak, waiting to be released and to grow if the right kinds of sociocultural opportunities and pressures arise. For example, developmental psychologist Deanna Kuhn and her colleagues (Kuhn, Shaw, & Felton, 1997) have documented in a number of studies, some observing young adolescents of limited education, that repeated interaction and discussion on a topic enhances the quality of reasoning about that topic. It appears that minimal guidance from adults is needed for young people to adopt a generally rational stance—although the special skills of "winning" an argument may need to be explicitly supported or taught. And analyzing much younger, elementary school children's conversations about books they had read, Anderson and colleagues (et. al., 1997) found logical argument structures in children's talk on subjects as diverse as ecology problems and the motives of characters in children's literature.

Even if there is universal human *potential* for rational discourse, however, special care is needed to create educative environments in which that potential is *nurtured* and emerges as a characteristic of whole social communities. In modern societies, the most likely place in which to develop the skills of reasoned discourse on a wide scale is the school, for that is where societies have an opportunity to influence the norms of interaction that will govern the discourse of all the diverse people now populating our countries.

Unfortunately, an occasional lesson—or even a semester's course in "critical thinking" or "logical reasoning"—cannot produce the habits and practices of reasoned discourse. Reasoned discourse is a *habit,* a *way of life.* It must be *socialized,* learned by living daily for many months and years in an environment that expects such behavior, supports it, and rewards it in overt and subtle ways. The only venue in which there is any hope of achieving such widespread socialization is the school. To develop the "turn of mind" that privileges reasonable argument, discursive practice will have to permeate students' school experience—every day and in all of the disciplines to which our schools are committed: science, mathematics, literature, history, and more.

The fact that each discipline has its own genre of talk (discussing poetry is not much like discussing mathematics) means that students can have opportunities to observe practice and join in several different kinds of discourse. Just as the math student must explain the solution to a problem, the student of history must clearly explain the actions that lead up to an event. Apprenticeship in the core disciplines of schooling provides the necessary structure to acquire these discourse-based reasoning abilities. The socialization of discourse skills requires that members of the school community—student, teachers, administrators—must make these skills commonplace in every classroom, not simply currency of the debate club.

This will take time, but it need not be time taken *away* from academic disciplines. To the contrary, considerable evidence now indicates that carefully designed and implemented talk-based pedagogy is likely the most powerful way to learn academic disciplines. Discussion-based classroom practices that combine rigorous tasks with carefully orchestrated, teacher-led discussion can support the growth of both disciplinary knowledge and the capacity to engage in reasoned discussion. Evidence is now accumulating that these ways of learning produce broad capacities that we label "growing the mind."

THE DISCURSIVE CLASSROOM

Over the past two decades, research has accumulated on how discussion methods are used in classrooms and why such discussion may support learning of important school subject matter as well as the process of reasoned participation. This research draws on sociocultural principles that emphasize the importance of social practices—in particular, the careful orchestration of talk and tasks in academic learning and in scholarly research. Much of this work has been done in the content areas of mathematics and science (Chapin, O'Connor, & Anderson, 2003; Resnick, Bill, & Lesgold, 1992; Yackel & Cobb, 1996), where students are expected not only to master a body of authoritative knowledge (algorithms, formulas, and symbolic tools as well as facts and accepted theories) but also to be able to reason with the ideas and tools of others. Sense making and scaffolded discussion, calling for particular forms of talk, are seen as the primary mechanisms for promoting deep understanding of complex concepts and robust reasoning. Multiple studies in mathematics, and to a lesser extent science learning, have demonstrated the role of certain kinds of structured talk for learning with understanding (Ball & Bass, 2003; Cobb, 2001; Delpit & Dowdy, 2002; Forman, Larreamendy-Joerns, Stein, & Brown, 1998; Goldenberg, 1992/3; Mercer, 2002; Michaels, Sohmer, & O'Connor, 2004; O'Connor, 2001; O'Connor, Hall, & Resnick, 2002; O'Connor & Michaels, 1996; Pontecorvo, 1993; Walqui &

Koelsch, 2006; Warren & Rosebery, 1996; Wells, 2001; Yackel & Cobb, 1996).

A number of studies suggests that this kind of classroom discourse leads to deeper engagement in the content under discussion and surprisingly elaborated, subject matter–specific reasoning by students who might not normally be considered able students (Chapin & O'Connor, 2004; Cobb, Boufi, McClain, & Whitenack, 1997; Lampert, 2001; Michaels, 2005; O'Connor, 1999; Resnick & Nelson-Le Gall, 1997). Opportunities for students to reflect and communicate about their mathematical work have been identified as essential for learning mathematics with understanding (Hiebert et al., 1997). During discussion, students can see how others approach a task and can gain insight into solution strategies and reasoning processes that they may not have considered. By engaging in whole-class, teacher-guided reflective discourse, students can explain their reasoning; make mathematical generalizations and connections between concepts, strategies, or representations; and benefit from the collective work of the class for a given lesson or task (Cobb et al., 1997). Some scholars have stressed the importance of participation in discussions for English language learners and other students with limited experience with and exposure to academic language (Adger, Snow, & Christian, 2002; August & Hakuta, 1998; Baugh, 1999; Cocking & Mestre, 1988; Heath 1983; Lee, 2001; Moll, Amanti, Neff, & Gonzalez, 1992; Moschkovich, 2000; Moses & Cobb, 2001; Walqui & Koelseh, 2006).

There are a number of other "success stories" in the literature on instructional change and school reform where similar kinds of discourse-intensive instruction in difficult "traditional" high-demand subject matter has produced unexpected results (Ball & Lampert, 1998; Boaler & Greeno, 2000; Chapin et al., 2003; Fischbein, Jehiam, & Cohen, 1995; Merenluoto & Lehtinen, 2002; Minstrell, 1989; Tsamir, 2003; Vosniadou, Baltas, & Vamvakoussi, 2007). But the effectiveness of discourse-intensive instruction depends significantly on the quality of the mathematical tasks used in instruction. A growing body of research indicates that tasks with high-level cognitive demands are important in improving students' performance on state and national tests of mathematical achievement (Fuson, Carroll, & Druek, 2000; Riordan & Noyce, 2001; Schoen, Fey, Hirsch, & Coxford, 1999), in improving students' understanding of important mathematical concepts (Ben-Chaim, Fey, Fitzgerald, Benedetto, & Miller, 1998; Huntley, Rasmussen, Villarubi, Sangtong, & Fey, 2000; Reys, Reys, Lappan, Holliday, & Wasman, 2003;

Thompson & Senk, 2001), and in improving students' abilities to reason, communicate, solve problems, and make mathematical connections.

From Cupcakes to Equations

Fifteen years ago, an exceptional teacher, Victoria Bill (Bill, Leer, Reams, & Resnick, 1992; Resnick, Bill, Lesgold, & Leer, 1991), demonstrated the potential of classroom discourse in primary grade mathematics instruction. Students were taught by a whole-class instruction method in which a problem was posed to the class and the teacher then led a structured discussion in which students explained why different routes to solution were mathematically correct. They then conjectured and tested what would happen if, for example, quantities were changed or different geometric shapes were involved. In these discussions, children were being socialized into a form of mathematical thinking in which multiple solutions and perspectives were recognized. The most obvious difference in this type of mathematics classroom from the standard initiation-response-evaluation (IRE) recitation is that the children's contributions, albeit short and almost always in response to the teacher's requests, are incorporated by the teacher into relatively extended lines of reasoning that may last across several teacher–student exchanges. These conversations may also contain overt challenges to children's responses—counterexamples—and children are expected to develop reasoned replies in response.

Bill's work was observed by a number of linguists and others interested in how to use talk to get students to reason with one another's ideas. O'Connor and Michaels (1993, 1996) developed their concept of "revoicing" as a critical strategy in such teaching, based in part on analyses of Bill's classroom.

Exhibit 7.2 contains a short excerpt of a portion of Bill's teaching. The children were mostly African Americans in an inner-city school. At the outset of a lesson, the teacher poses a practical problem—figuring out, *without counting the cupcakes one by one*, whether she and her daughters made enough cupcakes the night before for the whole second grade class.

Several cohorts of Bill's students, almost all lower-income minority children, showed dramatic increases in computational skill and significant increases in comprehension and problem solving on standardized tests. This success led the school system in which this math teaching took place to train teachers throughout the district in these methods—with

Exhibit 7.2

SHORT EXCERPT OF A PORTION OF BILL'S TEACHING

I made blueberry cupcakes last night. Lauren and Lisa arranged the cupcakes on the tray. Lauren said, "There are 3 rows of 7." Lisa said, "There are 7 rows of 3." Are there enough cupcakes for 2nd grade?

The teacher leads the discussion, showing a tray with cupcakes on it to two students on either side of a table and begins:

Teacher:	Well, okay, "there are 3 rows," so she was sitting over here, she said there's three rows and each row has seven.
	One two *(teacher points at each cupcake in a row as the class counts up to seven).*
Class:	three, four, five, six, seven
Teacher:	And Lisa was sitting over here and she said how many rows were there? *Teacher changes orientation of the tray)*
Teacher and Class:	one, two, three *(teacher points to each row as the class counts)*
Teacher:	Is that . . . Go ahead. Oh oh. Mark. She said they just changed the tray around. Is that all they did?
Class:	Yeah, yes.
Teacher:	You mean there's still, Mia, is there still going to be the same number of cupcakes on this tray?
Mia:	Yes
Teacher:	But seven, seven rows sounds like more. Rachael, What do you think? Seven, seven rows sounds like more than, I'm sorry, seven rows sound like more than (*teacher moves a tray of cupcakes to illustrate both Lauren and Lisa's representation of cupcake rows*).
Rachael:	Three rows.
Teacher:	Three rows. What do you think? What do you think? Sounds like more . . .

results less dramatic but still positive. This method of teaching developed more or less simultaneously in several American mathematics education projects such as QUASAR (Stein, Grover, & Henningsen, 1996; Stein, Smith, & Silver, 1999), Cognitively Guided Instruction (Carpenter, Fennema, Franke, Levi, & Empson 1999), and programs in Europe and Japan.

Existentialism

Another example, this one from a high school English class—again in an inner-city school—shows the range of academically rigorous discursive practice. Exhibit 7.3 is a transcript of a few minutes of discussion. In this discussion, high-school seniors engaged in an analysis of two very different types of literature: Jean Paul Sartre's *Existentialism* and James Baldwin's *Sonny's Blues.* The dialogue sample that appears in Exhibit 7.3 is

Exhibit 7.3

STUDENTS DISCUSSING SARTRE AND BALDWIN

Student 2:	Yeah, I agree with that, like even the last section, when um, he was, he was at Isabella's house, he was always on the piano, and Isabella had told Sonny's brother that it was like, not like living with a person, it was like living with sounds.
Student 1:	Yeah.
Student 2:	It was like he was separate from them. Like they had their lives and he had his own life. While they still provided for him, they still never really got close to him. Never really did anything with them. It was just like, him, and then it was them.
Student 3:	What do you have for your meaning of forlornness?
Student 2:	Mine? Uh, being alone.
Student 3:	I mean, well, in the Sartre essay it says, "When we speak of forlornness, a term Heidegger was fond of, we mean that, only that God does not exist and that we have to face all the consequences of this." So I mean like, I don't, I don't, I still don't understand how that truly shows forlornness.
Student 4:	My definition was kind of similar to that, I had, "the feeling of being alone in your responsibility, not having a higher power to blame."
Student 3:	So I don't understand how that shows forlornness. Like, the, the existentialist meaning of forlornness - I don't think it does.

(Continued)

Exhibit 7.3

STUDENTS DISCUSSING SARTRE AND BALDWIN *(Continued)*

Student 1:	What'd you say, what, what, what was the last part of your uh, definition?
Student 4:	" . . . not having a higher power . . . not having a higher power to blame."
Student 1:	Well I think this is, it's his feeling, he wanted to play, he wanted to be a musician. You know, it's, I think it was on him, I mean no one chose it for him, but, it was like, that kind of separated, like Will said, kind of separated himself from the rest of his family. And, like, the definition says, he doesn't have any, you know, higher being to blame, it, it was all his choice, like, you know, Sonny didn't, I mean his brother didn't want him to be a drummer. So it's like uh, you know, he wanted, he, this is what he feels, as a, I think this is like a deep feeling that he has, that he wanted to be a musician, so, it was like, it was no one else's fault but his own, you know? He's kind of on a . . . separate island, so to speak.
Student 3:	But was Sonny trying to blame it on anyone else, and then he like, I mean is he fleeing from forlornness, or is it for-, is it forlornness?
Student 2:	He's showing forlornness . . .
Student 1:	He's showing it, yeah.
Student 2:	Himself, cause like, he makes the decision. He's the, no one else knew about him cutting school to go down to, like, I don't know where he went, but he went down, and he was with other musicians, and he was playing down there, trying to hone his skills and trying to make a name for himself. And that was his responsibility, that was his decision, and he was a, he alone made those decisions, so I think he was just showing forlornness there, as well as anguish.
Student 3:	I mean I think it shows the definition, I mean the, excuse me, the uh, the dictionary meaning of forlornness, I don't, I don't, I still don't see how it really shows, you know what I'm saying, he's, he.

taken from a group of four students grappling with Sartre's definition of *forlornness* and attempting to decide whether Baldwin's Sonny is forlorn according to Sartre's and/or the dictionary definition. This discussion followed 4 weeks of teacher-led discussion modeling specific talk formats and behaviors, extensive note taking, and other preparatory work by students.

Although there were no formal leadership roles, it seems clear that Student 3 has taken it upon himself to keep the discussion focused on the *forlornness* question, pressing for definitions, asking for reference to texts, and using his puzzlement as a means of keeping the group focused. It is common for students to play this role, and they typically practice (as in this instance) strategies modeled earlier by the teacher.

ACCOUNTABLE TALK

The dialogues shown above are but examples of the kind of sophisticated discourse and reasoning moves students can develop in classrooms that provide the needed structure, time, and tools for learning. Over the past 15 years, our research has entailed intensive collaborations with teachers and students in many classroom contexts. We have confronted the challenges and limitations of contexts in which the discourse norms we seek are not initially shared by all members of the classroom community. We have worked to discover what it takes to lay the foundations for a discourse culture that includes veterans as well as newcomers, making the discourse norms and moves accessible to all. We call our work Accountable Talk®.

Accountable talk grows out of a Vygotskian theoretical framework (Wertsch, 1991) that emphasizes the "social formation of mind"—that is, the importance of social interaction in the development of individual mental processes. Similar ideas were also developed in seminal work by Dewey (1966) and George Herbert Mead (1967). In the accountable talk form of classroom interaction, the teacher poses a question that calls for a relatively elaborated response. As initial responses come forth, the teacher then presses a group of students to develop explanations, challenges, counterexamples, and questions. The process includes extended exchanges between teacher and student and among students and includes a variety of "talk moves," such as asking other students to explain what the first respondent has said, challenging students—sometimes via "revoicing" a student's contribution ("So let me see if I've got your idea right. Are you saying . . .?"), which makes the student's idea, reformulated by the teacher, available to the entire group.

Our work on accountable talk—across a wide range of classrooms and grade levels—suggests that its critical features fall under three broad dimensions: (a) accountability to the community, (b) accountability to standards of reasoning, and (c) accountability to knowledge. Students who learn school subject matter in classrooms guided by accountable talk standards are socialized into communities of practice in which respectful and grounded discussion, rather than noisy assertion or uncritical acceptance of the voice of authority, is the norm. The three aspects of accountable talk in combination are essential for the full development of student capacities and dispositions for reasoned civic participation (Michaels, O'Connor, Hall, & Resnick, 2002). We discuss each aspect separately in the following sections.

Accountability to the Learning Community

This is talk that attends seriously to and builds on the ideas of others; participants listen carefully to one another, build on each other's ideas, and ask each other questions aimed at clarifying or expanding a proposi-

Table 7.1

PROTOTYPICAL ACCOUNTABLE TALK MOVES

TALK MOVE ACTION	PROTOTYPICAL FORM
Revoicing.	"So let me see if I've got your thinking right. You're saying XXX?" (With time for students to accept or reject the teacher's formulation.)
Asking students to restate someone else's reasoning.	"Can you repeat what he just said in your own words?"
Asking students to apply their own reasoning to someone else's reasoning.	Do you agree or disagree and why?"
Prompting students for further participation.	"Would someone like to add on?"
Asking students to explicate their reasoning.	"Why do you think that?" or "How did you arrive at that answer?" or "Say more about that."
Challenge or counterexample.	"Is this always true?" "Can you think of any examples that would not work?"

tion. When talk is accountable to the community, participants listen to others and build their contributions in response to those of others. They make concessions and partial concessions ("*yes . . .but . . .*") and provide reasons when they disagree or agree with others. They may extend or elaborate someone else's argument, or ask someone for elaboration of an expressed idea (see Table 7.1).

This *community* facet of accountability seems to be the most straightforward and simple to implement in a classroom. Once introduced to the idea, teachers quickly find that a relatively small number of conversational moves seem to evoke the desired features of student talk.

The six most important talk moves and an example of each in its prototypical form follow.

When teachers regularly use these and similar conversational moves, it is typical that, a few weeks later, students can be heard using the following kinds of statements on their own:

> I disagree with Nelia, and I agree with Jamal.
> Um, that . . . can you repeat that question again?
> José, that gave me an idea. Um, what he said at first, that you have to turn them into fractions . . .
> I wanted to add something. She was probably trying to say . . .
> I agree now with Alex because . . .

In other words, students appropriate the accountability to community moves easily and create new norms of classroom discourse. These kinds of conversational norms and practices go a long way toward instantiating a culture of deliberation. However, it is very important to note that in order for the students to begin using these forms of talk, there have to be interesting and complex ideas to talk and argue about. Implicitly or explicitly, teachers who have implemented these discourse strategies have shifted away from simple questions and one-word answers and opened up the conversation to problems that support multiple positions or solution paths.

Once this kind of talk from students appears, another interesting thing happens. Teachers start to remark that they are amazed at what their students have to say. "I had no idea they were so smart," is a commonly heard remark from teachers new to accountable talk. "I was amazed to hear Leticia saying that; she's never talked before." "I was amazed by all the different ideas they came up with, and how they justified their ideas with evidence." It seems that simply opening up the conversation, with interesting and complex problems to support the talk along with a few key talk moves, gives teachers more access to the thinking, knowledge, and reasoning capabilities of their diverse students.

Accountability to Standards of Reasoning

This is talk that emphasizes logical connections and the drawing of reasonable conclusions. It is talk that involves explanation and self-correction. It often involves searching for premises rather than simply supporting or attacking conclusions. Research mentioned earlier in this paper (Resnick et al., 1993) suggests that adhering to standards of reasoning is something that people do quite naturally, although it is necessary to use tools of linguistic and logical analysis to detect the rationality of ordinary conversational discussions. Several other lines of research suggest that guided practice without direct instruction in reasoning standards or strategies can lead to improved interactive reasoning. Even very young children, in trying to understand and influence the world around them, can build arguments or question the premises of others' claims. These ideas may be undeveloped, incomplete, or even incorrect. But young children have far more to build on than was recognized in the past.

An example of a teacher-guided discussion in a kindergarten classroom provides a compelling example of the possibilities. As part of a unit called "Seeing Ourselves in Measurement," Ms. Martinez's kindergarteners were about to measure themselves to create a full-size height chart with a picture of each student at the appropriate height. Before they got started, Ms. Martinez said she had an important measurement question to ask them, and they had to come to a decision as an entire group. This vignette is adapted from Michaels and colleagues (Michaels, Shouse, & Schweingruber, 2008).

Ms. Martinez began by asking "Should we measure your heights with or without your shoes on? Sit down in your circle time spots and let's discuss this as scientists. Think about it first by yourself for a minute, and then let's talk."

Hands went up. "I have an idea." "I know." Ms. Martinez waited until many hands were up. Then she said, "You're all going to get a chance to give us your ideas. But you have to listen really, really hard to what everyone says, so we can come up with a good decision." Ms. Martinez called on Alexandra.

Alexandra: I think we should do it with our shoes off because some of our shoes are little and some are big or like high up. That wouldn't be fair.

Ms. Martinez: What do you mean by fair? Can you say a bit more about that?

Alexandra: You know. Someone might be taller because of their shoes but not really taller. That wouldn't be fair.

Ms. Martinez: Does anyone want to add on to what Alexandra said? Or does anyone disagree?

Ramon: (Ramon spoke Spanish at home and was just beginning to learn English.) I no agree. Shoes all the same. All like this big. (With that he measured the bottom of his shoe and held up two fingers.) It make no difference.

Ms. Martinez: So let me see if I got your idea right. Are you saying that since we all have shoes on and they're all about the same size, it adds the same amount to everyone's height and so it *would* be fair? Is that what you're thinking?

Ramon: Uh huh, and no stand on tippy toes. [Laughter]

Damani: I think take our shoes off because some shoes are taller. Look at your shoes! They're way bigger up. (He pointed to Ms. Martinez's shoes, which had 2-inch heels.) And mine are short and Lexi's are tall.

Damani pointed to his own shoes, which were slip-on sandals with flat heels, and then to Lexi, who was wearing shoes with thick rubber soles. By now several children had their legs in the air, showing off their shoes.

Ms. Martinez: Okay friends, we have a disagreement here. What are we going to do to make a decision? Alexandra's saying that it wouldn't be fair and Ramon is saying it wouldn't make any difference. Damani says it would make a difference. How should we decide?

Kataisha: We could line up our shoes and measure them and see if they're all the same. But you can see that some of them are not the same, so I don't think we really need to measure them all. Lexi's are really big and mine are not so big. That wouldn't be fair.

Ramon said he changed his mind. Now he thought no shoes would be better. In 10 minutes of discussion, the group had arrived at a consensus. Ms. Martinez was impressed. At first she had thought that a vote might settle things, but instead the students had used evidence and a shared sense of fairness. They were able to explain their reasons with evidence (the height of the heels) and challenge someone else's evidence with counterevidence. They were even able to propose a simple experiment (measure all the shoes) to evaluate a particular claim. They were able to hear each other out, agree and disagree, and even change their minds as new evidence was introduced. As kindergarteners, they were able to reason about the idea of a "fair test," which later in their education they would be able to extend to the idea of holding variables constant and the principle of a common point of origin in measurement.

Accountability to Knowledge

This brings us to the most complex of our three accountabilities—accountability to knowledge. Talk that is accountable to knowledge is based explicitly on facts, written texts, or other publicly accessible information. Speakers make an effort to get their facts right and make explicit the evidence behind their claims or explanations. They challenge each other when evidence is lacking or unavailable. When the content under discussion involves new or incompletely mastered knowledge, accountable discussion can uncover misunderstandings and misconceptions. A knowledgeable and skilled teacher is required to provide authoritative knowledge when necessary and to guide conversation toward academically correct concepts.

Of the three facets of accountability, accountability to knowledge is—perhaps surprisingly—the most difficult to achieve and the most contested. Some educators argue that the teaching and accumulation of facts is trivial, and teachers should not "tell" students an answer or teach them isolated factoids. Others say factual knowledge is foundational, and that before students can reason cogently they must acquire a great deal of factual information in any given domain.

In educational circles, getting the facts right and engaging discursively are often treated as if they were mutually exclusive. In the "curriculum wars," one group stresses accurate knowledge (to be acquired by direct instruction and practice) while the other emphasizes the processes of engagement regardless of "correct" facts. The dichotomy fails, however, under the lens of cognitive research on reasoning and knowledge acquisition (Bransford, Brown, & Cocking, 2000). Good reasoning, hence good discourse, depends on good knowledge. At the same time, the acquisition of good knowledge depends on active processing and good reasoning. Knowledge and reasoning develop best in tandem; neither precedes the other. Yet it is no easy task to orchestrate this interdependent development. Indeed, teaching good knowledge using discursive methods is perhaps pedagogy's greatest challenge.

In a development that would not have been predicted 20 years ago, the discipline in which discursive pedagogy has been most fully elaborated is mathematics. An international consensus has emerged on what high cognitive demand/high mathematical content in schools should look like. Two characteristics are central. First is the need for a task for students to work on that poses a genuine mathematical problem (Stein, Smith, Henningsen, & Silver, 2000). Second is the need for carefully structured talk by teachers who understand the important mathematical ideas embedded in the task. This structured talk combines accountability to mathematical knowledge with accountability to reasoning and to the community.

One final example, this time, in a middle-school mathematics classroom, illustrates how this combination works. The task appears in Figure 7.4; it calls for student to "mathematize" a situation—to define it in terms of variables and their relationships rather than to just find a particular answer. Exhibit 7.4 provides a transcript of part of the teacher-led discussion that occurred after small subgroups of students had been working to develop equations expressing the relationship in the problems. Now reconvened into a large group, the students take turns presenting and explaining their groups' equations and making connections between the visual representation of the building and their explanations. The teacher prompts the students to discuss the relationship between the number of cubes, the building number, and the equivalence between the equations. The teacher supports the students' efforts to build accountability in these ways: *to community:* incorporating different students' solutions in the problem-solving process; *to reasoning:* focusing students' attention on how different mathematical formulations can yield equivalent answers; and *to knowledge:* calling for mathematical language, including definitions of variables and equations.

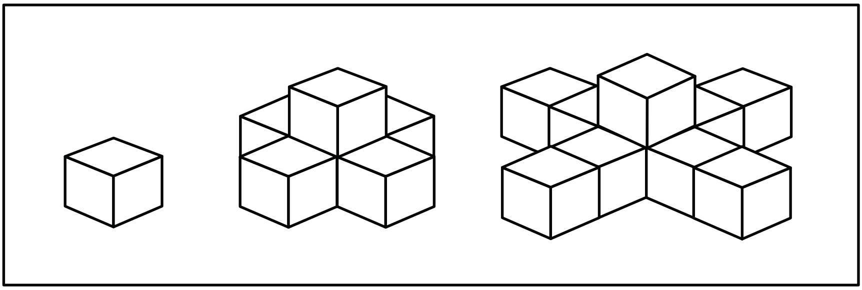


Figure 7.4 The Nth problem, adapted from "Counting Cubes," Connected Mathematics Project.

Exhibit 7.4

TRANSCRIPT OF STUDENTS DISCUSSING Nth BUILDING

Zach:	Okay, see Irina's (he points to work on board), says here, 1 + 5 (n–1)
	So, if you use the distributive property here, then it'd be, 1 + 5n – 5 Which is the same as, 5n – 4
Teacher:	Could you show that on the board? How that, how that is equivalent?
Zach:	(On the blackboard, Zach writes three equations)
	(a) 1 + 5 (n–1)
	(b) 1 + 5n – 5
	(c) –4 + 5n = 5n – 4
Teacher:	Okay . . . are those two things the same here? This left side (teacher points to –4 + 5n) of the equal sign, and the right? Is –4 + 5n = 5n – 4 (teacher points to 5n – 4)?
Students:	Yeah. Yes.
Teacher:	Okay, so we can say that, so I think what Zach is saying, is that, uh, Irina's, and Zach's group (equation) is identical?
Zach:	Yeah.
Teacher:	Okay, your group's is the same, Irina's form (points to other group's work), is the same as your group's—they're identical. What about this one? How does this one fit in, with that? Is there a mathematical equivalence there somehow? Yoshio, do you think you can show us that, or, explain it, or, how is it different, how is it the same?

TRANSCRIPT OF STUDENTS DISCUSSING Nth BUILDING *(Continued)*

Yoshio:	Our definition of n is different from theirs. For ours it's *n* equals the length of each arm. So, uh, the equation will change, will be different from the two.
Teacher:	So what you're telling me is that the definition of the variable is a very important idea in mathematics?
Students:	Yeah.
Teacher:	Okay, it makes the whole difference of . . . what the expression means?
Student 4:	Cause in there they're multiplying the, they're considering each arm, the, what we were considering plus the middle, and then they were going to subtract 4 middles, cause it would have 4 extra middles. And then, but what we were doing is we were just multiplying each of the arms without the middle, and then adding one middle. And it's really just the same thing. It's just, depends on how you think of it.

DOES ACCOUNTABLE TALK REALLY "BUILD THE MIND?"

Evidence is beginning to accumulate that discourse-intensive instruction for children does, in fact, support learning. Specifically, we have seen that accountable talk leads to the ability to engage in reasoning that is accountable to the community, to standards of reasoning, and to knowledge. However, a central question remains: Does this kind of instruction lead to generally improved cognitive performance? We are just at the beginning of being able to provide an empirically grounded answer to this question. Most efforts to teach using discursive methods that engage students in content-specific talk have not followed the students over long enough periods and into different enough settings to produce conclusive evidence. We are just now mounting a program of research that will use the capacities of advanced computer technology, including natural language recognition and machine learning, to be able to collect and analyze enough data of this kind to be able to provide the empirical evidence we need. There are two lines of research that suggest what we may be able to find. The first was that done by two British researchers and the second by one of us (O'Connor), using an approach similar to Accountable Talk.

Work by Adey and Shayer (2001) at King's College, London, is called the CASE intervention (Cognitive Acceleration through Science Education). For 2 years, 11- to 14-year-old students in eight British schools were given, within their science curriculum, special lessons intended to promote formal operational thinking. The students were prompted to articulate their thoughts about various "thinking science" activities. Three years after the end of the intervention program, students' achievement was tested by their results on British National examinations, taken at age 16. Although the intervention was set within the context of science learning, significant effects were found in mathematics and English, as well. In comparison with control classes, the effect sizes were 0.72 for mathematics and 0.69 for English.

Recent research conducted by two colleagues at Boston University (Chapin & O'Connor, 2004), has provided us with further evidence that discourse-based teaching can, in fact, "build the mind." Project Challenge was a 4-year intervention led by scholar/researchers at Boston University. Its purpose was to provide challenging mathematics education for potentially talented students in Chelsea, Massachusetts, the lowest-performing district in the state. Project Challenge served 400 Chelsea students, starting in fourth grade and following through to seventh grade. More than 70% of these students qualified for lunch aid, and over 60% spoke languages other than English at home.

The Project Challenge intervention included 1-hour class every day of high-demand mathematical tasks and structured teacher-led talk. Project Challenge teachers were trained to use a variety of academically productive talk moves and talk formats designed to press students to explicate their reasoning and build on one another's thinking. After 2 years, the proportion of students rated as showing a "high probability of giftedness in mathematics" on the Test of Mathematical Abilities rose from 4% to 41%. At the end of 2½ years, the class average on the California achievement test showed that Project Challenge students performed at high levels in computation, mathematical understanding, and problem solving. Most impressive of all, after 3 years of this kind of instruction, 82% of the Project Challenge students scored "Advanced" or "Proficient" on the Massachusetts state assessment—which is generally judged to be the most demanding in the United States. The state average of proficiency was 38% (see Figure 7.5). Finally, in a post hoc, quasi-controlled comparison of students who had been eligible for Project Challenge but were not selected, the differences between Project Challenge students and their matched controls was significant and effect sizes were large (1.8); (see Figure 7.6). Most surprising of all, there was transfer to English.

THE FUTURE OF EDUCATION THAT BUILDS THE MIND

Assuming that at least some of the studies now under way by various investigators show similar patterns of transfer and long-term retention as the two vanguard effort described above have, we will have to look for a theoretical explanation. How could participation in relatively brief and highly focused discussions produce such robust and long-lasting changes, including transfer? There has to be something quite general at work—it is apparently not just the specific mathematics or science knowledge that is learned; there is some kind of far transfer going on, the kind of elusive "holy grail" that learning and developmental psychologists have sought for generations. This effect, if it proves real, is particularly mysterious because its success seems to be rooted in detailed attention to very particular tasks and knowledge. The students and the discussion are not focused on general strategies of reasoning but rather on explaining in discipline-specific language, what certain expressions mean, why certain solutions to a problem work and others do not, and what an author intends when using a specific word.

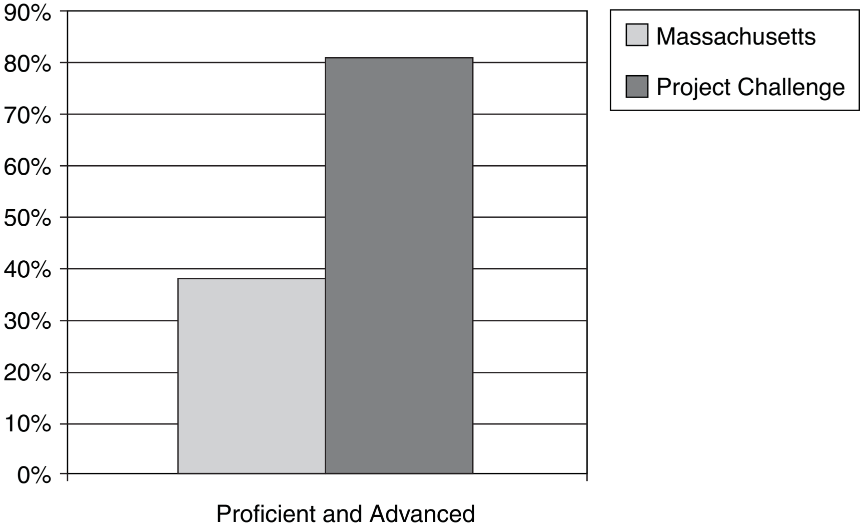


Figure 7.5 Project Challenge students' scores on the Massachusetts state test.

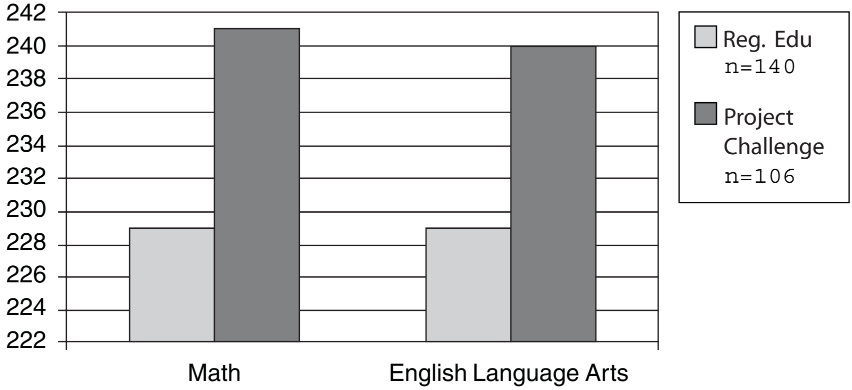


Figure 7.6 Project Challenge students compared with post-hoc matched control.

How might this disciplined particularity work to produce generally robust learning? There are several possibilities: A first is the development of a taste for explanation. Chi (2000) has shown that engaging in self-explanation as one reads a text produces better learning of the content and better retention. Our accountable talk classrooms embed this request for self-explanation in a social setting that may sustain this taste over time. A second possibility is that accountable talk classrooms provide children with language for conducting argumentation in ways that are socially acceptable. Yet a third possibility is that the experience of discussing a single problem or text over time provides an experience of "going deep" and persevering on a difficult task that some children might not experience regularly, at least not in the context of academic work. In a carefully orchestrated discussion, children of very different levels of knowledge and linguistic capacity can come away with the sense of being "good at" a kind of language and social participation that is associated with high achievement. In this way, participation over time in accountable talk environments may actually "socialize intelligence."

REFERENCES

Adey, P., & Shayer, M. (2001). *Thinking science.* London: Nelson Thormes.

Adger, C. T., Snow, C. E., & Christian, D. (Eds.). (2002). *What teachers need to know about language.* Washington, DC: Center for Applied Linguistics.

Anderson, R. C., Chinn, C., Chang J., Waggoner, M., & Yi, H. (1997). On the logical integrity of children's arguments. *Cognition and Instruction, 15,* 135–167.

August, D., & Hakuta, K. (Eds.). (1998). *Educating language minority children.* Washington, DC: National Academy Press.

Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group,* 3–14. Edmonton, AB: CMESG/GCEDM.

Ball, D. L., & Lampert, M. (1998). Multiples of evidence, time, and perspective: Revising the study of teaching and learning. In E. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 371, 398). San Francisco: Jossey-Bass.

Baugh, J. (1999). *Out of the mouths of slaves: African American language and educational malpractice.* Austin: The University of Texas Press.

Ben-Chaim, D., Fey, J. T., Fitzgerald, W. M., Benedetto, C., & Miller, J. (1998). Proportional reasoning among 7th grade students with different curricular experiences. *Educational Studies in Mathematics, 36*(3), 247–273.

Bill, V. L., Leer, M. N., Reams, L. E., & Resnick, L. B. (1992). From cupcakes to equations: The structure of discourse in a primary mathematics classroom. *Verbum, 15*(1), 63–85.

Blair, J. A., & Johnson, R. H. (1987). Argumentation as dialectical. *Argumentation, 1,* 41–56.

Boaler, J., & Greeno, J. G. (2000). Identity, agency and knowing in mathematical worlds. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 171–200). Westport, CT: Ablex.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999), *Children's mathematics: Cognitively guided instruction.* Portsmouth, NH: Heinemann.

Chapin, S. & O'Connor, M. C. (2004). Project Challenge: Identifying and developing talent in mathematics within low-income urban schools. *Boston University School of Education Research Report No. 1,* 1–6.

Chapin, S., O'Connor, C., & Anderson, N. (2003). *Classroom discussions: Using math talk to help students learn: Grades 1–6.* Sausalito, CA: Math Solutions Publications.

Chi, M. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah, NJ: Erlbaum.

Cobb, P. (2001). Supporting the improvement of learning and teaching in social and institutional context. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: 25 years of progress* (pp. 455–478). Mahwah, NJ: Erlbaum.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal of Research Into Mathematics Education, 28*(3), 258–277.

Cocking, R., & Mestre, J. (1988). *Linguistic and cultural influences on learning mathematics.* Hillsdale, NJ: Erlbaum.

Delpit, L, & Dowdy, J. K. (Eds.). (2002). *The skin that we speak: Thoughts on language and culture in the classroom.* New York: New Press.

Dewey, J. (1966). *Liberalism and social action*. New York: Putnam.

Fischbein, E., Jehiam, R., & Cohen, D. (1995). The concept of irrational numbers in high-school students and prospective teachers. *Educational Studies in Mathematics, 29,* 29–44.

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction, 8*(6), 527–548.

Fuson, K. C., Carroll, W. M., & Druek, J. V. (2000). Achievement results for second and third graders using standards-based curriculum Everyday Mathematics. *Journal for Research in Mathematics Education, 31*(3), 277–295.

Goldenberg, C. (1992/3). Instructional conversations: Promoting comprehension through discussion. *Reading Teacher, 46,* 316–326.

Grice, H. P. (1968). *Logic and conversation: The William James Lectures.* Cambridge, MA: Harvard University Press.

Habermas, J. (1990). Discourse ethics: Notes on a program of philosophical justification. In *Moral consciousness and communicative action* (pp. 43–115). Cambridge, MA: MIT Press.

Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms.* New York: Cambridge University Press.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., et al. (1997). *Making sense: Teaching and learning mathematics with understanding.* Portsmouth, NH: Heinemann.

Huntley, M. A., Rasmussen, C. L., Villarubi, R. S., Sangtong, J., & Fey, J. T. (2000). Effects of Standards-based mathematics education: A study of the Core-Plus Mathematics project algebra and function strand. *Journal for Research in Mathematics Education, 31*(3), 328–361.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction, 15*(3), 287–315.

Lampert, M. (2001). *Teaching problems and the problems of teaching.* New Haven, CT: Yale University Press.

Lee, C. (2001). Is October Brown Chinese? A cultural modeling activity system for underachieving students. *American Educational Research Journal, 38*(1), 97–141.

Mead, G. H. (1967). Mind, self, and society: From the standpoint of a social behaviorist. In C. W. Morris (Ed.), *Mind, self, and society: From the standpoint of a social behaviorist.* Chicago: University of Chicago Press.

Mercer, N. (2002). Developing dialogues. In G. Wells & G. Claxton (Eds.), *Learning for life in the 21st century: Sociocultural perspectives on the future of education* (pp. 141–153). Oxford: Blackwell.

Merenluoto, K., & Lehtinen, E. (2002). Conceptual change in mathematics: Understanding the real numbers. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change. Issues in theory and practice* (pp. 233–258). Dordrecht: Kluwer.

Michaels, S. (2005). Can the intellectual affordances of working-class storytelling be leveraged in school? *Human Development, 48*, 136–145.

Michaels, S., O'Connor, C., Hall, M., & Resnick, L. (2002). *Talk: Classroom conversation that works* (CD-ROM set). Pittsburgh, PA: University of Pittsburgh.

Michaels, S., Shouse, A., & Schweingruber, H. (2008). *Reading, set, science!: Putting research to work in K-8 science classrooms.* Washington, DC: National Academies Press.

Michaels, S., Sohmer, R., & O'Connor, M. C. (2004). Classroom discourse. In H. Ammon, N. Dittmar, K. Mattheier, & P. Trudgill (Eds.), *Sociolinguistics: An international handbook of the science of language and society* (2nd ed., pp. 2351–2366). New York: de Gruyter.

Minstrell, J. (1989). Teaching science for understanding. In L. B. Resnick & L. E. Lopfer (Eds.), *Toward the thinking curriculum: Current cognitive research (1989 yearbook of the Association for Supervision and Curriculum Development)*; (pp. 131–149). Alexandria, VA: Association for Supervision and Curriculum Development.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31,* 132–141.

Moschkovich, J. N. (2000). Learning mathematics in two languages: Moving from obstacles to resources. In W. Secada (Ed.), *Changing the faces of mathematics: Vol. 1. Perspectives on multiculturalism and gender equity* (pp. 85–93). Reston, VA: NCTM.

Moses, R. P., & Cobb, C. E. (2001). *Radical equations: Math literacy and civil rights.* Boston: Beacon.

O'Connor, C., Hall, M., & Resnick, L. (2002). *Talk: Classroom conversation that works* (CD-ROM set). Pittsburgh, PA: University of Pittsburgh.

O'Connor, M. C. (1999). Language socialization in the mathematics classroom: Discourse practices and mathematical thinking. In M. Lampert & M. Blunk (Eds.), *Talking mathematics* (pp. 17–55). Cambridge, UK: Cambridge University Press.

O'Connor, M. C. (2001). "Can any fraction be turned into a decimal?" A case study of a mathematical group discussion. *Educational Studies in Mathematics, 46,* 143–185.

O'Connor, M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education Quarterly, 24,* 318–335.

O'Connor, M. C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning and schooling* (pp. 63–103). Cambridge, UK: Cambridge University Press.

Pontecorvo, C. (1993). Forms of discourse and shared thinking. *Cognition and Instruction, 11*(3 & 4), 189–196.

Resnick, L. B., Bill, V., & Lesgold, S. (1992). Developing thinking abilities in arithmetic class. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-Piagetian theories of cognitive development: Implications and applications for education* (pp. 210–230). London: Routledge.

Resnick, L. B., Bill, V., Lesgold, S., & Leer, M. (1991). Thinking in arithmetic class. In B. Means, C. Chelemer, & M. S. Knapp (Eds.), *Teaching advanced skills to at-risk students: Views from research and practice* (pp. 27–53). San Francisco: Jossey-Bass.

Resnick, L. B, Hall, M. W., & Fellows of the Institute for Learning. (2001). *The principles of learning: Study tools for educators* (CD-ROM). Pittsburgh, PA: Institute for Learning, Learning Research and Development Center, University of Pittsburgh.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.). (1991). *Perspectives on socially shared cognition*. Washington, DC: American Psychological Association.

Resnick, L. B., & Nelson-Le Gall, S. (1997). Socializing intelligence. In L. Smith, J. Dockrell, & P. Tomlinson (Eds.), *Piaget, Vygotsky and beyond* (pp. 145–158). London: Routledge.

Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction, 11,* 347–364.

Reys, R. E., Reys, B. J., Lappan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standards-based middle grades mathematics curriculum materials on student achievement. *Journal for Research in Mathematics Education, 34*(1), 74–95.

Riordan, J. E., & Noyce, P. E. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32*(4), 368–398.

Schoen, H. L., Fey, J. T., Hirsch, C. R., & Coxford, A. F. (1999). Issues and opinions in the math wars. *Phi Delta Kappan, 80*(6), 444–453.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455–488.

Stein, M. K., Smith, M. S., Henningsen, M., & Silver, E. A. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development.* New York: Teachers College Press.

Stein, M. K., Smith, M. S., & Silver, E. A. (1999). The development of professional developers: Learning to assist teachers in new settings in new ways. *Harvard Educational Review, 69*(3), 237–269.

Thompson, D. R., & Senk, S. L. (2001). The effects of curriculum on achievement in second-year algebra: The example of the University of Chicago School Mathematics Project. *Journal for Research in Mathematics Education, 32*(1), 58–84.

Toulmin, S. E. (1958). *The uses of argument.* Cambridge, UK: Cambridge University Press.

Tsamir, P. (2003). Primary intuitions and instruction: The case of actual infinity. *Research in Collegiate Mathematics Education, 12,* 79–96.

van Eemeren, F. R., & Grootendorst, R. (1988). Rules for argumentation in dialogues. *Argumentation, 2*(4), 499–510.

Vosniadou S., Baltas, A., & Vamvakoussi, X. (2007). *Reframing the conceptual change approach in learning and instruction.* Oxford: Elsevier.

Voss, J. F., Perkins, D. N., & Segal, J. W. (Eds.). (1991). *Informal reasoning and education.* Hillsdale, NJ: Erlbaum.

Walqui, A., & Koelsch, N. (2006). *Scaffolding academic uses of English: Accelerating the achievement of secondary school English learners.* San Francisco: Jossey-Bass.

Walton, D., & Krabbe, E. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning.* Albany: State University of New York Press.

Warren, B., & Rosebery, A. (1996). "This question is just too, too easy": Perspectives from the classroom on accountability in science. In L. Schauble & R. Glasser (Eds.), *Innovations in learning: New environments for education* (pp. 97–125). Hillsdale, NJ: Erlbaum.

Wells, G. (Ed.). (2001). *Action, talk, and text: Learning and teaching through inquiry.* New York: Teachers College Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action.* Cambridge, MA: Harvard University Press.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.

8 A Framework for Critical Thinking, Rational Thinking, and Intelligence

KEITH E. STANOVICH AND PAULA J. STANOVICH

Critical thinking is highly valued in educational writings if not in practice. Despite a substantial literature on the subject, for many years the area of critical thinking was notorious for its conceptual difficulties. For example, years ago Cuban (1984) lamented that "defining thinking skills, reasoning, critical thought, and problem solving is troublesome to both social scientists and practitioners. Troublesome is a polite word; the area is a conceptual swamp" (p. 676). There has been some progress in elucidating the concept of critical thinking since the time of Cuban's statement, but we shall argue here that educational theory is on the verge of an even more stunning conceptual advance in the area of critical thinking. Education is beginning to understand the critical thinking concept by relating it to the constructs of intelligence and rational thought. In fact, modern cognitive science provides a coherent framework for understanding the relation between critical thinking, intelligence, and rational thought.

THE FOUNDATIONAL SKILLS OF CRITICAL THINKING

In the critical thinking literature, the ability to evaluate evidence and arguments independently of one's prior beliefs and opinions is a skill that is strongly emphasized (Baron, 2008; Dole & Sinatra, 1998; Ennis, 1987, 1996; Kuhn, 2005; Lipman, 1991; Paul, 1984, 1987; Ritchhart & Perkins, 2005; Sternberg, 1997, 2001, 2003; Wade & Tavris, 1993). The disposition toward such unbiased reasoning is almost universally viewed as a characteristic of good thinking. For example, Norris and Ennis (1989) argue that one fundamentally important characteristic of critical thinking is the disposition to "reason from starting points with which we disagree without letting the disagreement interfere with reasoning" (p. 12). Zechmeister and Johnson (1992) list as one characteristic of the critical thinker the ability to "accept statements as true even when they don't agree with one's own position" (p. 6). Similarly, Nickerson (1987) stresses that critical thinking entails the ability to recognize "the fallibility of one's own opinions, the probability of bias in those opinions, and the danger of differentially weighting evidence according to personal preferences" (p. 30). The growing literature on informal reasoning likewise emphasizes the importance of detaching one's own beliefs from the process of argument evaluation (Baron, 1995; Klaczynski & Lavallee, 2005; Kuhn, 2001, 2007; Kuhn & Udell, 2007; Macpherson & Stanovich, 2007; Toplak & Stanovich, 2003; Voss, Perkins, & Segal, 1991).

The emphasis on unbiasedness in reasoning has led many theorists to highlight the importance of decontextualization as the foundational skill of critical thinking thought (see Paul, 1984, 1987; Siegel, 1988). For example, Kelley (1990) argues that "the ability to step back from our train of thought . . . . is a virtue because it is the only way to check the results of our thinking, the only way to avoid jumping to conclusions, the only way to stay in touch with the facts" (p. 6). Neimark (1987) lumps the concepts of decentering and decontextualizing under the umbrella term *detachment*. She terms one component of detachment *depersonalizing*: being able to adopt perspectives other than one's own. This aspect of detachment is closely analogous to Piaget's (1926) concept of decentration. Neimark's (1987) second component of detachment—detaching from context—involves breaking the bounds of situational constraint and local context. It is reminiscent of Donaldson's (1978, 1993) concept of disembedding:

> If the intellectual powers are to develop, the child must gain a measure of control over his own thinking and he cannot control it while he remains unaware of it. The attaining of this control means prising thought out of its primitive unconscious embeddedness in the immediacies of living in the world and interacting with other human beings. It means learning to move

> beyond the bounds of human sense. It is on this movement that all the higher intellectual skills depend. (Donaldson, 1978, p. 123)

Neimark (1987) emphasizes how associations built up over time will tend to activate a decision for us automatically and unconsciously if we are not reflective and cannot detach from situational cues. The danger of response patterns that are determined too strongly by overlearned cues is a repeated theme in the heuristics and biases literature of cognitive science (e.g., Arkes, 1991; Evans, 2003, 2006, 2007, 2008; Kahneman, 2003; Kahneman & Frederick, 2002; Stanovich, 2003, 2004, 2009; Wilson & Brekke, 1994). For example, Baron (1994) argues that many departures from consequentialism in decision making are due to inappropriate generalizations. For example, the act–omission distinction is hypothesized to arise because harmful acts are usually more intentional than harmful omissions; but this distinction continues to be made even when there is no difference in intention. In short, to act in consequentialist fashion, the features of the actual context (intention, etc.) must be abstracted and compared componentially. Such decontextualizing cognitive habits represent one line of defense against overlearned associations that might trigger nonnormative responses.

Many tasks in the heuristics and biases branch of the reasoning literature might be said to involve some type of decontextualization skill (Kahneman, 2003; Stanovich, 2003). Tasks are designed to see whether reasoning processes can operate independently of interfering context (world knowledge, prior opinion, vivid examples). One example of such a task is the laboratory paradigm that has been used to study belief bias in syllogistic reasoning. The stimuli in the task put the validity of a syllogism and the facts expressed in the conclusion of the syllogism in conflict. For example, the syllogism "All flowers have petals, roses have petals, therefore roses are flowers" is invalid despite the seeming "rightness" of the conclusion. The inability to decouple prior knowledge from reasoning processes has been termed the *belief bias effect* (Evans, Barston, & Pollard, 1983; Evans, Newstead, Allen, & Pollard, 1994), and it has been the subject of extensive study in the cognitive science literature (De Neys, 2006; Evans, 2002; Evans & Curtis-Holmes, 2005; Evans & Feeney, 2004; Garnham & Oakhill, 2005; Klauer, Musch, & Naumer, 2000).

Belief bias has also been revealed in paradigms where participants must evaluate the quality of empirical evidence in a manner not contaminated by their prior opinions on the issue in question. In several studies, Klaczynski and colleagues (Klaczynski, 1997; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000) presented participants with flawed hypothetical experiments leading to conclusions that were either consistent or inconsistent with prior positions and opinions. Participants then critiqued the flaws in the experiments (which were most often badly flawed). Participants found many more flaws when the experiment's conclusions were inconsistent with their prior opinions than when the experiment's conclusions were consistent with their prior opinions and beliefs (see also Macpherson & Stanovich, 2007).

CRITICAL THINKING IN THE SERVICE OF RATIONAL THOUGHT

Like the study of wisdom (Sternberg, 2001, 2003; Sternberg & Jordan, 2005), the study of critical thinking is a normative/evaluative endeavor. Specifically, if one's goal is to *aid* people in their thinking, then it is essential that one have some way of *evaluating* thinking. For example, in the current educational literature, teachers are constantly exhorted to "teach children how to think," or to foster "critical thinking" and "creative problem solving." However, the problem here is that "thinking" is not a domain of knowledge. As Baron (1993) notes,

> We teach Latin or calculus because students do not already know how to speak Latin or find integrals. But, by any reasonable description of thinking, students already know how to think, and the problem is that they do not do it as effectively as they might. (p. 199)

Thus the admonition to educators to "teach thinking skills" and foster "critical thinking" contains implicit evaluative assumptions. The children *already* think. Educators are charged with getting them to think *better* (Adams, 1993). This, of course, implies a normative model of what we mean by better thinking (Baron, 1993, 2008).

A somewhat analogous issue arises when thinking dispositions are discussed in the educational literature of critical thinking. Why do we want people to think in an actively open-minded fashion? Why do we want to foster multiplist and evaluative thinking (Kuhn, 1993, 2001, 2005; Kuhn & Udell, 2007) rather than absolutist thinking? Why do we want people to be reflective? It can be argued that the superordinate goal we are actually trying to foster is that of rationality (Stanovich, 2004, 2009). That is, much of what educators are ultimately concerned about is rational thought in both the epistemic sense and the practical sense. We value certain thinking dispositions because we think that they will at least aid in bringing belief in line with the world and in achieving our goals. By a parallel argument, we could equally well claim that the superordinate goal is to educate for wisdom (Sternberg, 2001, 2002, 2003).

We can see that it is rationality, and not critical thinking per se, that is the ultimate goal of education by conducting some simple thought experiments or imaginative hypotheticals. For example, we could imagine a person with excellent epistemic rationality (his or her degree of confidence in propositions being well calibrated to the available evidence relevant to the proposition) and optimal practical rationality (the person optimally satisfies desires) who was *not* actively open-minded—that is, who was not a good critical thinker under standard assumptions. Of course we would still want to mold such an individual's dispositions in the direction of open-mindedness for the sake of society as a whole. But the essential point for the present discussion is that, from a purely *individual* perspective, we would now be hard pressed to find reasons why we would *want* to change such a person's thinking dispositions—whatever they were—if they had led to rational thought and action in the past.

In short, a large part of the rationale for educational interventions to change thinking dispositions derives from a tacit assumption that actively open-minded critical-thinking dispositions make the individual a more rational person—or as Sternberg (2001, 2005) argues, a wiser, less foolish person. Thus, the normative justification for fostering critical thought is that it is the foundation of rational thought. The thinking dispositions associated with critical thinking must be fostered because they make students more rational. Our view is consistent with that of many other theorists who have moved toward conceptualizing critical thinking as a subspecies of rational thinking or at least as closely related to rational thinking (Kuhn, 2005; Moshman, 2004, 2005, 2010; Reyna, 2004; Siegel, 1988, 1997).

The grounding of critical thinking within the concept of rationality in this manner has many conceptual advantages. First, the concept of rationality is deeply intertwined with the data and theory of modern cognitive science (see LeBoeuf & Shafir, 2005; Over, 2004; Samuels & Stich, 2004; Stanovich, 2004, 2009) in a way that the concept of critical thinking is not. Additionally, as we demonstrate below, theory in cognitive science differentiates rationality from intelligence and explains why rationality and intelligence often dissociate. This finding and its explanation confirm the long-standing belief in education that intelligence does not guarantee critical thinking.

RATIONAL THOUGHT AND ITS OPERATIONALIZATIONS IN COGNITIVE SCIENCE

Cognitive scientists recognize two types of rationality: instrumental and epistemic. The simplest definition of instrumental rationality is behaving in the world so that you get exactly what you most want, given the resources (physical and mental) available to you. Somewhat more technically, we could characterize instrumental rationality as the optimization of the individual's goal fulfillment. Economists and cognitive scientists have refined the notion of optimization of goal fulfillment into the technical notion of expected utility. The model of rational judgment used by decision scientists is one in which a person chooses options based on which option has the largest expected utility (see Baron, 2008; Dawes, 1998; Hastie & Dawes, 2001; Wu, Zhang, & Gonzalez, 2004).

The other aspect of rationality studied by cognitive scientists is termed *epistemic rationality*. This aspect of rationality concerns how well beliefs map onto the actual structure of the world. Epistemic rationality is sometimes called theoretical rationality or evidential rationality (see Audi, 1993, 2001; Foley, 1987; Harman, 1995; Manktelow, 2004; Over, 2004). Instrumental and epistemic rationality are related. The aspect of beliefs that enter into instrumental calculations (that is, tacit calculations) are the probabilities of states of affairs in the world.

One of the fundamental advances in the history of modern decision science was the demonstration that if people's preferences follow certain patterns (the so-called axioms of choice—things like transitivity and freedom from certain kinds of context effects), they are behaving as if they were maximizing utility—they are acting to get what they most want (Edwards, 1954; Jeffrey, 1983; Luce & Raiffa, 1957; Savage, 1954; von Neumann & Morgenstern, 1944). This is what makes people's degrees of rationality measurable by the experimental methods of cognitive science. Although it is difficult to assess utility directly, it is much easier to assess whether one of the axioms of rational choice is being violated. This has been the logic of the research program on heuristics and biases inaugurated in the much-cited studies of Kahneman and Tversky (1972, 1973, 1979; Tversky & Kahneman, 1974, 1981, 1983, 1986).

Researchers in the tradition of heuristics and biases have demonstrated in a host of empirical studies that people violate many of the strictures of rationality and that the magnitude of these violations can be measured experimentally. For example, people display confirmation bias, test hypotheses inefficiently, display preference inconsistencies, do not properly calibrate degrees of belief, overproject their own opinions onto others, combine probabilities incoherently, and allow prior knowledge to become implicated in deductive reasoning (for summaries of the large literature, see Baron, 2008; Evans, 1989, 2007; Gilovich, Griffin, & Kahneman, 2002; Kahneman & Tversky, 2000; Shafir & LeBoeuf, 2002; Stanovich, 1999, 2004, 2009). These are caused by many well-known cognitive biases: base-rate neglect, framing effects, representativeness biases, anchoring biases, availability bias, outcome bias, and vividness effects, to name just a few. Degrees of rationality can be assessed in terms of the number and severity of such cognitive biases that individuals display. Failure to display a bias becomes a measure of rational thought.

In an attempt to understand how these various errors in rational thinking originate, investigators working in the tradition of heuristics and biases have been inexorably drawn to dual-process models of cognitive architecture. Recently, in an attempt to extend these models into the domain of individual differences, Stanovich (2009) has proposed a triprocess distinction that both explains errors in heuristics and biases tasks and, even more importantly, elucidates the relation between rationality and intelligence.

DUAL-PROCESS MODELS OF COGNITION

Virtually all attempts to classify heuristics and biases tasks end up utilizing a dual-process framework because most of the tasks in the literature on heuristics and biases were deliberately designed to pit a heuristically triggered response against a normative response generated by the analytic system. As Kahneman (2000) notes, "Tversky and I always thought of the heuristics and biases approach as a two-process theory" (p. 682). Since Kahneman and Tversky launched the heuristics and biases approach in the 1970s, a wealth of evidence has accumulated in support of the dual-process approach. Evidence from cognitive neuroscience and cognitive psychology converges on the conclusion that mental functioning can be characterized by two different types of cognition having somewhat different functions and different strengths and weaknesses (Brainerd & Reyna, 2001; Evans, 2003, 2008, 2009; Evans & Over, 1996, 2004; Feldman Barrett, Tugade, & Engle, 2004; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Kahneman & Frederick, 2002, 2005; Lieberman, 2003; McClure, Laibson, Loewenstein, & Cohen, 2004; Metcalfe & Mischel, 1999; Sloman, 1996, 2002; Stanovich, 1999, 2004).

There are many such theories (over 20 dual-process theories are presented in a table in Stanovich, 2004) and they have some subtle differences, but they are similar in that all distinguish autonomous from nonautonomous processing. The two types of processing were termed systems in earlier writings, but theorists have been moving toward more atheoretical characterizations; we therefore follow Evans (2009) in using the terms *type 1* and *type 2 processing*.

The defining feature of type 1 processing is its autonomy. Type 1 processes are termed autonomous because (a) their execution is rapid, (b) their execution is mandatory when the triggering stimuli are encountered, (c) they do not put a heavy load on central processing capacity (i.e., they do not require conscious attention), (d) they do not depend on input from high-level control systems, and (e) they can operate in parallel without interfering with each other or with type 2 processing. Type 1 processing would include behavioral regulation by the emotions, the encapsulated modules for solving specific adaptive problems that have been posited by evolutionary psychologists, processes of implicit learning, and the automatic firing of overlearned associations (see Evans, 2007, 2008; Stanovich, 2004, 2009).

Type 2 processing contrasts with type 1 processing on each of the critical properties that define the latter. Type 2 processing is relatively slow and computationally expensive—it is the focus of our awareness. Many type 1 processes can operate at once in parallel, but only one (or a very few) type 2 thoughts can be executing at once—type 2 processing is thus serial processing. Type 2 processing is often language-based.

One of the most critical functions of type 2 processing is to override type 1 processing. All of the different kinds of type 1 processing (processes of emotional regulation, Darwinian modules, associative and implicit learning processes) can produce responses that are irrational in a particular context if not overridden. In order to override type 1 processing, type 2 processing must display at least two (possibly related) capabilities. One is the capability of interrupting type 1 processing and suppressing its response tendencies. Type 2 processing thus involves inhibitory mechanisms of the type that have been the focus of recent work on executive functioning (Hasher, Lustig, & Zacks, 2007; Miyake, Friedman, Emerson, & Witzki, 2000; Zelazo, 2004).

But the ability to suppress type 1 processing gets the job only half done. Suppressing one response is not helpful unless a better response is available to substitute for it. Where do these better responses come from? One answer is that they come from processes of hypothetical reasoning and cognitive simulation that are unique to type 2 processing (Evans, 2007; Evans & Over, 2004; Kahneman & Tversky, 1982; Nichols & Stich, 2003). When we reason hypothetically, we create temporary models of the world and test out actions (or alternative causes) in that simulated world. In order to reason hypothetically we must, however, have one critical cognitive capability—the ability to distinguish our representations of the real world from representations of imaginary situations. For example, in considering an alternative goal state different from the one we currently have, we must be able to represent our current goal and the alternative goal and to keep straight which is which. Likewise, we need to be able to differentiate the representation of an action about to be taken from representations of potential *alternative* actions we are considering. But the latter must not infect the former while the mental simulation is being carried out. Thus, several years ago in a much-cited article, Leslie (1987) modeled pretense by positing a so-called secondary representation (see Perner, 1991) that was a copy of the primary representation but that was decoupled from the world so that it could be manipulated—that is, be a mechanism for simulation. The important issue for our purposes is that decoupling secondary representations from the world and then maintaining the decoupling while simulation is carried out is a type 2 processing operation. It is computationally taxing and greatly restricts the ability to do any other type 2 operation. In fact, decoupling operations might well be a major contributor to a distinctive type 2 property—its seriality.

Cognitive decoupling must take place when an individual engages in a simulation of alternative worlds in order to solve a problem. Problem-solving tasks that necessitate fully disjunctive reasoning (see Johnson-Laird, 2006; Shafir, 1994) provide examples of the situations that require fully decoupled simulation. Fully disjunctive reasoning involves considering all possible states of the world in deciding among options or in choosing a problem solution in a reasoning task. Consider the following problem, taken from the work of Levesque (1986, 1989) and studied by our research group (see Toplak & Stanovich, 2002): Jack is looking at

Anne but Anne is looking at George. Jack is married but George is not. Is a married person looking at an unmarried person?

A. Yes
B. No
C. Cannot be determined

The vast majority of people answer C (cannot be determined) when in fact the correct answer to this problem is A (yes). To answer correctly, both possibilities for Anne's marital status (married and unmarried) must be considered to determine whether a conclusion can be drawn. If Anne is married, then the answer is yes because she would be looking at George, who is unmarried. If Anne is not married, then the answer is still yes because Jack, who is married, would be looking at Anne. Considering all the possibilities (the fully disjunctive reasoning strategy) reveals that a married person is looking at an unmarried person whether Anne is married or not. The fact that the problem does not *reveal* whether Anne is married suggests to people that nothing can be determined. Many people make the easiest (incorrect) inference from the information given and do not proceed with the more difficult (but correct) inference that follows from fully disjunctive reasoning.

Not all type 2 processing represents fully explicit cognitive simulation, however. Or, to put it another way: all hypothetical thinking involves type 2 processing (Evans & Over, 2004) but not all type 2 processing involves hypothetical thinking. What has been termed serial associative cognition (Stanovich, 2009) represents this latter category. It can be understood by considering a discussion of the four-card selection task in a theoretical paper on dual processes by Evans (2006; see also Evans & Over, 2004). In Wason's (1966) four-card selection task, the participant is told the following:

> Each of the boxes below represents a card lying on a table. Each one of the cards has a letter on one side and a number on the other side. Here is a rule: If a card has a vowel on its letter side, then it has an even number on its number side. As you can see, two of the cards are letter-side up, and two of the cards are number-side up. Your task is to decide which card or cards must be turned over in order to find out whether the rule is true or false. Indicate which cards must be turned over.

The participant chooses from four cards labeled K, A, 8, and 5 (corresponding to not-P, P, Q, and not-Q). The correct answer is to pick the

A and the 5 (P and not-Q), but the most common answer is to pick the A and 8 (P and Q)—the so-called matching response.

Evans (2006) points out that the previous emphasis on the matching bias evident in the task (Evans, 1972, 1998; Evans & Lynch, 1973) might have led some investigators to infer that type 2 processing does not occur. In fact, matching bias might be viewed as just one of several such suggestions in the literature that much thinking during the task is type 1 processing (see Hardman, 1998; Margolis, 1987; Stanovich & West, 1998a; Tweney & Yachanin, 1985). In contrast, however, Evans (2006) presents evidence indicating that type 2 processing may be going on during the task—even on the part of the majority who do not give the normatively correct response but instead give the PQ response.

In fact, type 2 processing is occurring in this task, but it is not full-blown cognitive simulation of alternative world models. It is thinking of a shallower type—serial associative cognition. Serial associative cognition is not rapid and parallel, such as type 1 processes, but is nonetheless rather inflexibly locked into an associative mode that takes as its starting point a model of the world that is *given* to the subject. For example, Evans and Over (2004) note that in the studies of verbal protocols, when participants made an incorrect choice, they referred to the hidden sides of the cards they were going to pick, but referred only to verification when they did so. Thus, the evidence suggests that people accept the rule as given, assume it is true, and simply describe how they would go about verifying it. The fact that they refer to hidden sides does not mean that they have constructed any alternative model of the situation beyond what was given to them by the experimenter and their own assumption that the rule is true. They then reason from this single focal model—systematically generating associations from this focal model but never constructing another model of the situation. This is why the central characteristic of serial associative cognition is that it displays a *focal bias.*

One way in which to contextualize the idea of focal bias is as the second stage in a framework for thinking about human information processing that dates to the mid 1970s—the idea of humans as cognitive misers (Dawes, 1976; Taylor, 1981; Tversky & Kahneman, 1974). There are, in fact, two aspects of cognitive miserliness. Dual-process theory has heretofore highlighted only rule 1 of the cognitive miser: default to type 1 processing whenever possible. But defaulting to type 1 processing is not always possible—particularly in novel situations where there are no stimuli available to trigger domain-specific evolutionary modules. Type 2 processing procedures will be necessary, but a cognitive miser default is operating even there. Rule 2 of the cognitive miser is that when type 1 processing will not yield a solution, default to serial associative cognition with a focal bias (*not* fully decoupled cognitive simulation).

So the focal model that will dominate processing—the only model that serial associative cognition deals with—is the most easily constructed model. The most easily constructed model: tends to represent only one state of affairs; it accepts what is directly presented and models what is presented as true; it ignores moderating factors—probably because taking account of those factors would necessitate modeling several alternative worlds and this is just what a focal processing allows us to avoid. And finally, given the voluminous literature in cognitive science on belief bias and the informal reasoning literature on myside bias, the easiest models to represent clearly appear to be those closest to what a person already believes and has modeled previously (e.g., Evans & Feeney, 2004; Stanovich & West, 2007, 2008a). Thus, serial associative cognition is defined by its reliance on a single focal model that triggers all subsequent thought.

THREE KINDS OF MINDS

In 1996 philosopher Daniel Dennett wrote a book about how aspects of the human mind were like the minds of other animals and how other aspects were not. He titled the book *Kinds of Minds* to suggest that within the brain of humans are control systems of very different types—different kinds of minds. We are going to make here a distinction between aspects of type 2 processing that can be introduced through a Dennett-type example of different types of explanation. Imagine two different stories involving a woman walking on a cliff. The stories are all sad—the woman dies in each. The purpose of this exercise is to get us to think about how we explain the death in each story. In incident A, the woman is walking on a cliffside by the ocean and goes to step on a large rock, but the rock is not a rock at all. Instead, it is actually the side of a crevice and she falls down the crevice and dies. In incident B, the woman attempts suicide by jumping off a cliff and dies when she is crushed on the rocks below.

In both cases, at the most basic level, when we ask ourselves for an explanation of why the woman died, we might say that the answer is the same. The same laws of physics that are in operation in incident A (the gravitational laws that describe why the woman will be crushed upon impact) are also operative in incident B. However, we feel that the laws of gravity and force somehow do not provide a complete explanation of what has happened in incident B. This feeling is correct. The examples each call for a different level of explanation if we wish to zero in on the *essential* cause of death.

In analyzing incident A, a psychologist would be prone to say that in processing a stimulus (the crevice that looked somewhat like a rock) the woman's information processing system malfunctioned—sending the wrong information to response decision mechanisms, which then resulted in a disastrous motor response. Cognitive scientists refer to this level of analysis as the algorithmic level (Anderson, 1990; Marr, 1982; Stanovich, 1999). In the realm of machine intelligence, this would be the level of the instructions in the abstract computer language used to program the computer (FORTRAN, COBOL, etc.). The cognitive psychologist works largely at this level by showing that human performance can be explained by positing certain information processing mechanisms in the brain (input coding mechanisms, perceptual registration mechanisms, storage systems of short- and long-term memory, etc.). For example, a simple letter pronunciation task might entail encoding the letter, storing it in short-term memory, and comparing it with information stored in long-term memory and, if a match occurs, making a response decision and executing a motor response. In the case of the woman in incident A, the algorithmic level is the right level to explain her unfortunate demise. Her perceptual registration and classification mechanisms malfunctioned by providing incorrect information to response decision mechanisms, causing her to step into the crevice.

Incident B, on the other hand, does not involve such an information processing error at the algorithmic level. The woman's perceptual apparatus accurately recognized the edge of the cliff and her motor command centers accurately programmed her body to jump off the cliff. The computational processes posited at the algorithmic level of analysis executed quite perfectly. No error at this level of analysis explains why the woman is dead in incident B. Instead, this woman died because of her overall goals and how these goals interacted with her beliefs about the world in which she lived.

In the spirit of Dennett's book, Stanovich (2009) termed the part of the mind that carries out type 1 processing the *autonomous mind*. Different kinds of type 2 processing are defined by incidents A and B in the imaginary scenarios. In the terms of Stanovich (2009), the woman in incident A had a problem with the algorithmic mind and the woman in incident B had a problem with the reflective mind. This terminology captures the fact that we turn to an analysis of goals, desires, and beliefs to understand a case such as B. The algorithmic level provides an incomplete explanation of behavior in cases like incident B because it provides an information processing explanation of how the brain is carrying out a particular task (in this case, jumping off of a cliff) but no explanation of *why* the brain is carrying out this particular task. We turn to the level of the reflective mind, where we ask questions about the *goals* of the system's computations (*what* the system is attempting to compute and *why*). In short, the reflective mind is concerned with the goals of the system, beliefs relevant to those goals, and the choice of action that is optimal given the system's goals and beliefs. It is only at the level of the reflective mind that issues of rationality come into play. Importantly, the algorithmic mind can be evaluated in terms of efficiency but not rationality.

This concern for the efficiency of information processing as opposed to its rationality is mirrored in the status of intelligence tests. They are measures of efficiency but not rationality—a point made clear by considering a distinction that is very old in the field of psychometrics. Psychometricians have long distinguished typical performance situations from optimal (sometimes termed maximal) performance situations (see Ackerman, 1994, 1996; Ackerman & Heggestad, 1997; Ackerman & Kanfer, 2004; see also, Cronbach, 1949; Matthews, Zeidner, & Roberts, 2002). Typical performance situations are unconstrained in that no overt instructions to maximize performance are given and the task interpretation is determined to some extent by the participant. The goals to be pursued in the task are left somewhat open. The issue is what a person would typically do in such a situation, given few constraints. Typical performance measures are measures of the reflective mind—they assess in part goal prioritization and epistemic regulation. In contrast, optimal performance situations are those where the task interpretation is determined externally. The person performing the task is instructed to maximize performance and is told how to do so. Thus, optimal performance measures examine questions of efficiency of goal pursuit—they capture the processing efficiency of the algorithmic mind. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of critical or rational thinking are often assessed under typical performance conditions.

The difference between the algorithmic mind and the reflective mind is captured in another well-established distinction in the measurement of individual differences—the distinction between cognitive ability and thinking dispositions. The former are, as just mentioned, measures of the efficiency of the algorithmic mind. The latter travel under a variety of names in psychology—thinking dispositions or cognitive styles being the two most popular. Many thinking dispositions concern beliefs, belief structure and, importantly, attitudes toward forming and changing beliefs. Other thinking dispositions that have been identified concern a person's goals and goal hierarchy. Examples of some thinking dispositions that have been investigated by psychologists are: actively open-minded thinking, need for cognition (the tendency to think a lot), consideration of future consequences, need for closure, superstitious thinking, and dogmatism (Cacioppo, Petty, & Feinstein 1996; Kruglanski & Webster, 1996; Norris & Ennis, 1989; Schommer-Aikins, 2004; Stanovich, 1999, 2009; Sternberg, 2003; Sternberg & Grigorenko, 1997; Strathman, Gleicher, Boninger, & Scott Edwards, 1994).

The literature on these types of thinking dispositions is vast and our purpose is not to review that literature here. It is only necessary to note that the types of cognitive propensities that these thinking disposition measures reflect are the tendency to collect information before making up one's mind, to seek various points of view before coming to a conclusion, to think extensively about a problem before responding, to calibrate the degree of strength of one's opinion to the degree of evidence available, to think about future consequences before taking action, to explicitly weigh pluses and minuses of situations before making a decision, and to seek nuance and avoid absolutism. In short, individual differences in thinking dispositions include assessing variation in people's goal management, epistemic values, and epistemic self-regulation—differences in the operation of reflective mind. They are all psychological characteristics that underpin rational thought and action.

The cognitive abilities assessed on intelligence tests are not of this type. They are not about high-level personal goals and their regulation, the tendency to change beliefs in the face of contrary evidence, or how knowledge acquisition is internally regulated when not externally directed. People have indeed come up with *definitions* of intelligence that encompass such things. Theorists often define intelligence in ways that encompass rational action and belief; but nevertheless *the actual measures of intelligence in use assess only algorithmic-level cognitive capacity.* No current intelligence test that is even moderately used in practice assesses rational thought or behavior.

The algorithmic mind, assessed on actual IQ tests, is relevant in determining what happened in the case A above, but it does not provide sufficient explanation of case B. To understand what happened in case B, we need to know about more than the woman's processes of memory and speed of pattern recognition. We need to know what her goals were and what she believed about the world. And one of the most pressing things we want to know about the woman in case B is whether there was some sense in her jumping off the cliff. We do not want to know whether she threw herself off with the greatest efficiency possible (an algorithmic-level question). We want to know whether it was *rational* for her to jump.

A TRIPARTITE MODEL OF MIND AND INDIVIDUAL DIFFERENCES

We have now bifurcated the notion of type 2 processing into two different things—the reflective mind and the algorithmic mind. Previous dual-process views tended to ignore individual differences and hence to miss critical differences in type 2 processing. Figure 8.1 represents the classification of individual differences in the tripartite view. The broken horizontal line represents the location of the key distinction in older dual-process views. The figure identifies variation in fluid intelligence with individual differences in the efficiency of processing of the algorithmic mind. In contrast, thinking dispositions index individual differences in the reflective mind. The reflective and algorithmic minds are characterized by continuous individual differences. Continuous individual differences in the autonomous mind are few. Disruptions to the autonomous mind often reflect damage to cognitive modules, which results in very discontinuous cognitive dysfunction such as autism or the agnosias and alexias (Anderson, 2005; Bermudez, 2001; Murphy & Stich, 2000).

Figure 8.1 highlights an important sense in which rationality is a more encompassing construct than intelligence. To be rational, a person must have well-calibrated beliefs and must act appropriately on those beliefs to achieve goals—both properties of the reflective mind. The person must, of course, have the algorithmic-level machinery that enables him or her to carry out the actions and to process the environment in a way that allows the correct beliefs to be fixed and the correct actions to be taken. Thus individual differences in rational thought and action can arise because of individual differences in intelligence (the algorithmic mind) or because of individual differences in thinking dispositions (the reflective mind). To put it simply, the concept of ratio-

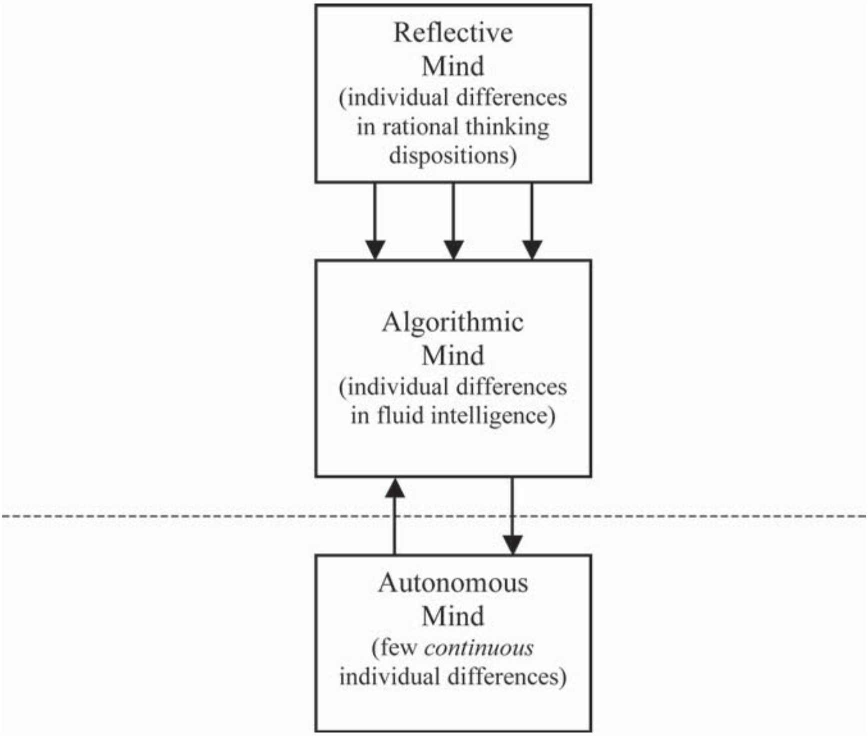


Figure 8.1 Individual differences in the tripartite model.

nality encompasses two things—thinking dispositions of the reflective mind and algorithmic-level efficiency—whereas the concept of intelligence, at least as it is commonly operationalized, is largely confined to algorithmic-level efficiency.

The conceptualization in Figure 8.1 has two great advantages. First, it conceptualizes intelligence in terms of what intelligence tests actually measure. That is, all current tests assess various aspects of algorithmic efficiency. But that is all they assess. None attempt to measure directly an aspect of epistemic or instrumental rationality, nor do they examine any thinking dispositions that relate to rationality. It seems perverse to define intelligence as including rationality when no existing IQ test measures any such thing!

The best-known indicators of cognitive functioning—intelligence and cognitive ability tests—do not assess a critical aspect of thinking, the ability to think rationally. To think rationally means adopting appropriate goals, taking the appropriate action given one's goals and beliefs, and holding beliefs that are commensurate with available evidence. Standard intelligence tests do not assess such functions (Perkins, 1995, 2002; Stanovich, 2002, 2009; Sternberg, 2003, 2006). For example, although intelligence tests do assess the ability to focus on an immediate goal in the face of distraction, they do not assess at all whether a person has the tendency to develop goals that are rational in the first place. Likewise, intelligence tests provide good measures of how well a person can hold beliefs in short-term memory and manipulate those beliefs, but they do not assess at all whether a person has the tendency to *form* beliefs rationally when presented with evidence. And again, similarly, intelligence tests give good measures of how efficiently a person processes information that has been provided, but they do not at all assess whether the person is a *critical assessor* of information as it is gathered in the natural environment.

It is clear from Figure 8.1 why rationality and intelligence can become dissociated. As long as variation in thinking dispositions is not perfectly correlated with fluid intelligence, there is the statistical possibility of dissociations between rationality and intelligence. Substantial empirical evidence indicates that individual differences in thinking dispositions and intelligence are far from perfectly correlated. Many different studies involving thousands of subjects (e.g., Ackerman & Heggestad, 1997; Austin & Deary, 2002; Baron, 1982; Bates & Shieles, 2003; Cacioppo et al., 1996; Eysenck, 1994; Goff & Ackerman, 1992; Kanazawa, 2004; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Zeidner & Matthews, 2000) have indicated that measures of intelligence display only moderate to weak correlations (usually less than .30) with some thinking dispositions (e.g., actively open-minded thinking, need for cognition) and near zero correlations with others (e.g., conscientiousness, curiosity, diligence).

Other important evidence supports the conceptual distinction made here between algorithmic cognitive capacity and thinking dispositions. For example, across a variety of tasks from the heuristics and biases literature, it has consistently found that rational thinking dispositions will predict variance in these tasks after the effects of general intelligence have been controlled (Bruine de Bruin, Parker, & Fischhoff, 2007; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Kokis et al., 2002; Macpherson & Stanovich, 2007; Newstead, Handley, Harley, Wright, & Farrelly, 2004; Parker &

Fischhoff, 2005; Sá & Stanovich, 2001; Stanovich & West, 1997, 1998c, 2000; Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007; Toplak & Stanovich, 2002).

Measures of thinking dispositions tell us about the individual's goals and epistemic values—and they index broad tendencies of pragmatic and epistemic self-regulation at a high level of cognitive control. The empirical studies cited indicate that these different types of cognitive predictors are tapping separable variance, and the reason that this is to be expected is because cognitive capacity measures such as intelligence and thinking dispositions map on to different levels in the tripartite model.

The functions of the different levels of control are illustrated more completely in Figure 8.2. There, it is clear that the override capacity itself is a property of the algorithmic mind; it is indicated by the arrow labeled A. However, previous dual-process theories have tended to ignore the higher-level cognitive function that initiates the override function in the first place. This is a dispositional property of the reflective mind that is related to rationality. In the model in Figure 8.2, it is represented by arrow B, which represents, in machine intelligence terms, the call to the algorithmic mind to override the type 1 response by taking it offline. This is a different mental function than the override function itself (arrow A), and we have presented evidence indicating that the two functions are indexed by different types of individual differences: the ability to sustain the inhibition of the type 1 response is indexed by measures of fluid intelligence, and the tendency to initiate override operations is indexed by thinking dispositions such as reflectiveness and need for cognition.

Figure 8.2 represents another aspect of cognition somewhat neglected by previous dual-process theories. Specifically, the override function has loomed large in dual-process theory but less so the simulation process that computes the alternative response making the override worthwhile. Figure 8.2 explicitly represents the simulation function as well as the fact that the call to initiate simulation originates in the reflective mind. The decoupling operation (indicated by arrow C) itself is carried out by the algorithmic mind and the call to initiate simulation (indicated by arrow D) by the reflective mind. Again, two distinct types of individual differences are associated with the initiation call and the decoupling operator—specifically, rational thinking dispositions with the former and fluid intelligence with the latter.

The model in Figure 8.2 defines a third critical function for the algorithmic mind in addition to type 1 processing override and enabling

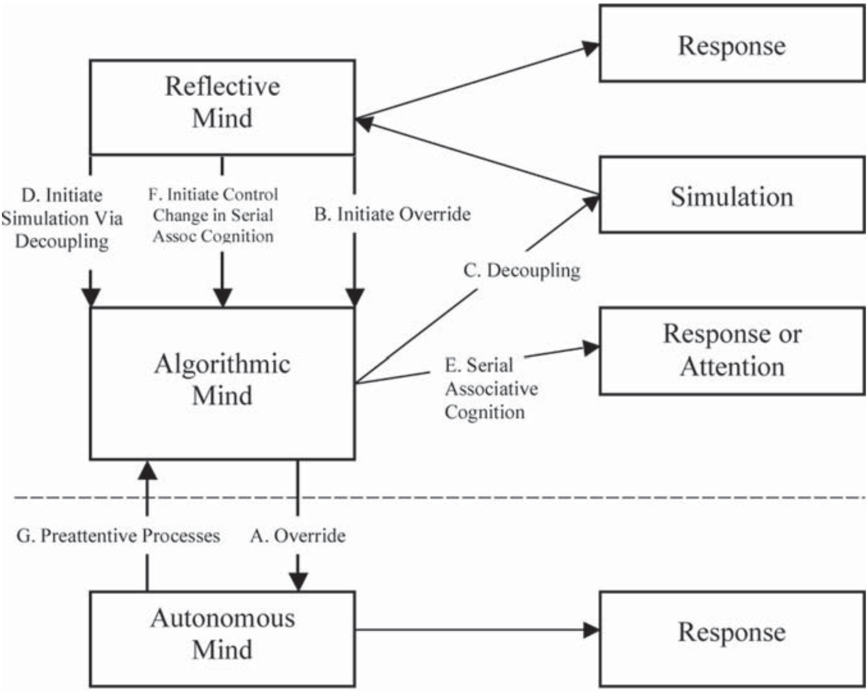


Figure 8.2 A more complete model of the tripartite structure.

simulation via decoupling. This third function is to sustain the serial associative cognition discussed previously (arrow labeled E). This function is there to remind us that not all type 2 processing involves strongly decoupled cognitive simulation. There are types of slow, serial cognition that do not involve simulating alternative worlds and exploring them exhaustively. The figure thus identifies a third function of the reflective mind—initiating an interrupt of serial associative cognition (arrow F). This interrupt signal alters the next step in a serial associative sequence that would otherwise direct thought. This interrupt signal might have a variety of outcomes. It might stop serial associative cognition altogether in order to initiate a comprehensive simulation (arrow C). Alternatively, it might start a new serial associative chain (arrow E) from a different starting point by altering the temporary focal model that is the source of a new associative chain. Finally, the algorithmic mind receive inputs from the computations of the autonomous mind (arrow G) via so-called preattentive processes (Evans, 2006, 2007, 2008, 2009).

THE IMPORTANCE OF MINDWARE

The term *mindware* was coined by David Perkins (1995) to refer to the rules, knowledge, procedures, and strategies that a person can retrieve from memory in order to aid decision making and problem solving. Perkins uses the term to stress the analogy to software in the brain/computer analogy. Each of the levels in the tripartite model of mind has to access knowledge to carry out its operations, as illustrated in Figure 8.3. As the figure indicates, the reflective mind accesses not only general knowledge structures but, importantly, also the person's opinions, beliefs, and reflectively acquired goal structure. The algorithmic mind accesses microstrategies for cognitive operations and production system rules for sequencing behaviors and thoughts. Finally, the autonomous mind accesses not only evolutionarily compiled encapsulated knowledge bases but also information that has become tightly compiled and available to the autonomous mind because of overlearning and practice.

It is important to note that Figure 8.3 displays the knowledge bases that are *unique* to each mind. Algorithmic- and reflective-level processes also receive inputs from the computations of the autonomous mind (see arrow G in Figure 8.2). The mindware available for retrieval, particularly that available to the reflective mind, is in part the product of past learning experiences. And here we have a direct link to the Cattell/Horn/Carroll theory of intelligence (Carroll, 1993; Cattell, 1963, 1998; Horn & Cattell, 1967), sometimes termed the fluid/crystallized (or Gf/Gc) theory. The theory posits that tests of mental ability tap a small number of broad factors, of which two are dominant. Fluid intelligence (Gf) reflects reasoning abilities operating across a variety of domains, including novel ones. It is measured by tests of abstract thinking such as figural analogies, Raven matrices, and series completion. Crystallized intelligence (Gc) reflects declarative knowledge acquired from acculturated learning experiences. It is measured by vocabulary tasks, verbal comprehension, and general knowledge measures. Ackerman (1996) discusses how the two dominant factors in the CHC theory reflect a long history of considering two aspects of intelligence: intelligence as process (Gf) and intelligence as knowledge (Gc).

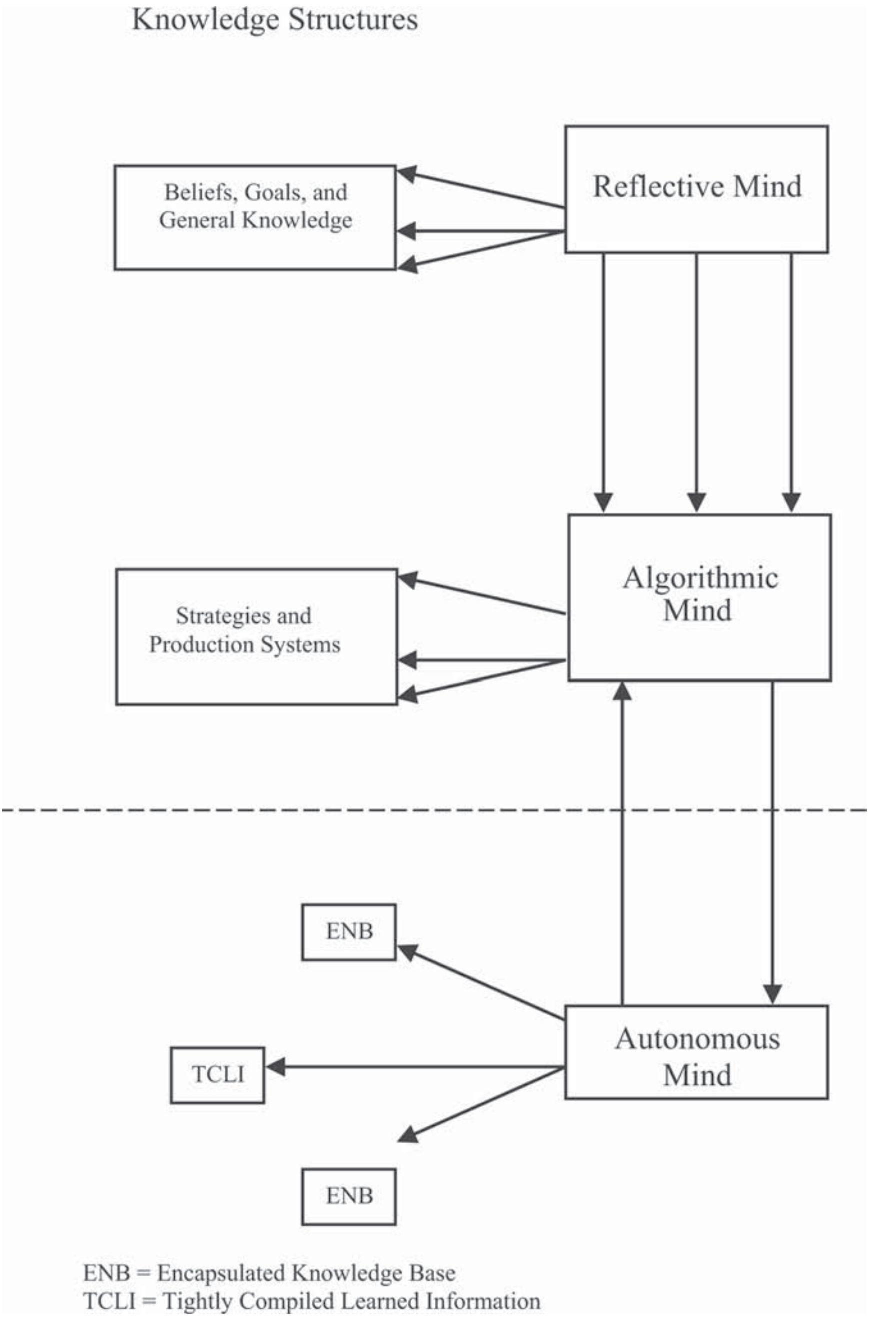


Figure 8.3 Knowledge structures in the tripartite model.

The knowledge structures available for retrieval by the reflective mind represent Gc (intelligence as knowledge). Recall that Gf (intelligence as process), is already represented in the Figure 8.2. It is the general computational power of the algorithmic mind—importantly exemplified by the ability to sustain cognitive decoupling.

Because the Gf/Gc theory is one of the more comprehensive theories of intelligence available that has extensive scientific validation, it is thus important to see how both of its major components miss critical aspects of rational thought. Gf will, of course, have some relation to rationality because it indexes the computational power of the algorithmic mind to sustain decoupling. Because override and simulation are important operations for rational thought, Gf will definitely facilitate rational action in some situations. Nevertheless, the tendency to initiate override (arrow B in Figure 8.2) and to initiate simulation activities (arrow D in Figure 8.2) are both aspects of the reflective mind unassessed by intelligence tests, so the tests will miss these components of rationality.

The situation with respect to Gc is a little different. It is true that much of the mindware of rational thought would be classified as crystallized intelligence in the abstract. But is it the kind of crystallized knowledge that is specifically assessed on the tests? The answer is no. The mindware of rational thought is somewhat specialized mindware (it clusters in the domains of probabilistic reasoning, causal reasoning, and scientific reasoning; see Stanovich, 2009). In contrast, the crystallized knowledge assessed on IQ tests is deliberately designed to be nonspecialized. The designers of the tests, in order to make sure the sampling of Gc is fair and unbiased, explicitly attempt to *broadly* sample vocabulary, verbal comprehension domains, and general knowledge. The broad sampling ensures unbiasedness in the test, but it inevitably means that the specific knowledge bases critical to rationality will go unassessed. In short, Gc, as traditionally measured, does not assess individual differences in rationality, and Gf will do so only indirectly and to a mild extent.

THE REQUIREMENTS OF RATIONAL THINKING

With this discussion of mindware, we have established that rationality requires three different classes of mental characteristic. First, algorithmic-level cognitive capacity is needed in order that override and simulation activities can be sustained. Second, the reflective mind must be characterized by the tendency to initiate the override of suboptimal responses generated by the autonomous mind and to initiate simulation activities that will result in a better response. Finally, the mindware that allows the computation of rational responses needs to be available and accessible during simulation activities. Intelligence tests assess only the first of these three characteristics that determine rational thought and action. As measures of rational thinking, they are radically incomplete.

Problems in rational thinking arise when the cognitive capacity is insufficient to sustain autonomous system override, when the necessity of override is not recognized, or when simulation processes do not have access to the mindware necessary for the synthesis of a better response. The source of these problems, and their relation to intelligence, helps to explain one data trend that has been uncovered (Stanovich & West, 2007, 2008a)—that many rational thinking problems are markedly independent of cognitive ability. Ever since Charles Spearman inaugurated the modern period of intelligence research in 1904, what he then termed positive manifold has been the ubiquitous finding: that intelligence indicators have correlated with a plethora of cognitive abilities and thinking skills that are almost too large to enumerate. This is why, among psychologists and among the lay public alike, assessments of intelligence are taken to be the sine qua non of good thinking. Critics of these instruments often point out that IQ tests fail to assess many mental traits outside of the cognitive domain, but these critiques concede too much. The tests miss critical thinking processes that are themselves cognitive—the numerous components of rational thinking.

Although the tests fail to assess rational thinking directly, it could be argued that the processes that are tapped by IQ tests largely overlap with variation in rational thinking ability. It is just this conjecture that research has contradicted. Consider the Levesque (1986, 1989) problem (e.g., "Jack is looking at Anne but Anne is looking at George") discussed above. The subjects who answer this problem correctly are no higher in intelligence than those who do not, at least in a sample of university students (see Toplak & Stanovich, 2002).

Most people can carry out fully disjunctive reasoning when they are explicitly *told* that it is necessary. But it is also true that most do not automatically do so. We might expect high-IQ individuals to excel at disjunctive reasoning when they know it is required for successful task performance. But the higher-IQ people in our sample were only slightly more likely to *spontaneously* adopt this type of processing in situations that do not explicitly require it. Note that the instructions in Levesque's

Anne problem do not cue the subject to engage in fully disjunctive reasoning. If *told* to reason through all of the alternatives, the subjects of higher intelligence probably would have done so more efficiently. However, without that instruction, they defaulted to computationally simple cognition in solving problems—they were cognitive misers like everyone else (see Stanovich, 2009). Intelligence and the tendency toward *spontaneous* disjunctive reasoning can be quite unrelated.

This tendency to process information incompletely has been a major theme throughout the past 30 years of research in psychology and cognitive science (Dawes, 1976; Taylor, 1981; Tversky & Kahneman, 1974). For example, Kahneman and Frederick (2002) have shown how people engage in attribute substitution—the substitution of an easy-to-evaluate characteristic for a harder one even if the easier one is less accurate. For example, the cognitive miser will substitute the less effortful attributes of vividness or salience for the more effortful retrieval of relevant facts. But when we are evaluating important risks—such as the risk of certain activities and environments for our children—we do not want to substitute vividness for careful thought about the situation. In such situations, we want to employ type 2 override processing to block the attribute substitution of the cognitive miser.

A simple example of miserly processing is discussed by Kahneman and Frederick (2002). They describe a simple experiment in which people were asked to consider the following puzzle: "A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?"

Many people offer the response that first comes to mind—10¢–without thinking further and realizing that this cannot be right. The bat would then have to cost $1.10 and the total cost would be $1.20 rather than the required $1.10. People often do not think deeply enough to make this simple correction though, and many students at very selective universities will answer incorrectly and move on to the next problem before realizing that their shallow processing has led them to make an error. Frederick (2005) has found that large numbers of highly selected students at MIT, Princeton, and Harvard, when given this and other similar problems, are cognitive misers like the rest of us. The correlation between intelligence and a set of similar items is quite modest, in the range of .40 to .50 (Gilhooly & Murphy, 2005).

Many other biases of the cognitive miser show correlations no greater than those shown in the Frederick bat-and-ball problem. In fact, some cognitive biases are almost totally dissociated from intelligence. Myside bias, for example, is virtually independent of intelligence (Macpherson & Stanovich, 2007; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007, 2008a, 2008b; Toplak & Stanovich, 2003). Individuals with higher IQs in a university sample are no less likely to process information from an egocentric perspective than are individuals with relatively lower IQs.

Irrational behavior can occur not just because of miserly processing tendencies but also because the right mindware (cognitive rules, strategies, knowledge, and belief systems) is not available to use in decision making. We would expect to see a correlation with intelligence here because mindware gaps most often arise from lack of education or experience. Nevertheless, while it is true that more intelligent individuals learn more things than less intelligent individuals, much knowledge (and many thinking dispositions) relevant to rationality are picked up rather late in life. Explicit teaching of this mindware is not uniform in the school curriculum at any level. That such principles are taught very inconsistently means that some intelligent people may fail to learn these important aspects of critical thinking. Correlations with cognitive ability have been found to be roughly (in absolute magnitude) in the range of .20 to .35 for probabilistic reasoning tasks and scientific reasoning tasks measuring a variety of rational principles (Bruine de Bruin, et al., 2007; Kokis et al., 2002; Parker & Fischhoff, 2005; Sá, West, & Stanovich, 1999; Stanovich & West, 1997, 1998b, 1998c, 1998d, 1999, 2000; Toplak & Stanovich, 2002). This is again a magnitude of correlation that allows for substantial discrepancies between intelligence and rationality. Intelligence is thus no inoculation against many of the sources of irrational thought. None of these sources of rational thought are directly assessed on intelligence tests, and the processes that *are* tapped by IQ tests are not highly overlapping with the processes and knowledge that explain variation in rational thinking ability.

Because the tasks used in this research are so various, we summarize some of this evidence by presenting Tables 8.1 and 8.2. Table 8.1 presents a sampling of rational thinking tasks—each task illustrating an important principle of rational thought—that have shown virtually no relation with intelligence in university samples. Table 8.2 presents a selection of effects and biases that show correlations in the .20 to .35 range.

Rationality is a multifarious concept—not a single mental quality. It requires various thinking dispositions that act to trump a variety of miserly information processing tendencies. It depends on the presence of various knowledge bases related to probabilistic thinking and scientific thinking. It depends on avoiding contaminated mindware that fosters

Table 8.1

TASKS THAT FAIL TO SHOW ASSOCIATIONS WITH COGNITIVE ABILITY
Noncausal base-rate usage (Stanovich & West, 1998c, 1999, 2008)
Conjunction fallacy between subjects (Stanovich & West, 2008)
Framing between subjects (Stanovich & West, 2008)
Anchoring effect (Stanovich & West, 2008)
Evaluability less is more effect (Stanovich & West, 2008)
Proportion dominance effect (Stanovich & West, 2008)
Sunk cost effect (Stanovich & West, 2008; Parker & Fischhoff, 2005)
Risk/benefit confounding (Stanovich & West, 2008)
Omission bias (Stanovich & West, 2008)
Perspective bias (Stanovich & West, 2008)
Certainty effect (Stanovich & West, 2008)
WTP/WTA difference (Stanovich & West, 2008)
My-side bias between and within S (Stanovich & West, 2007, 2008)
Newcomb's problem (Stanovich & West, 1999; Toplak & Stanovich, 2002)

Table 8.2

TASKS THAT SHOW .20–35 ASSOCIATIONS WITH COGNITIVE ABILITY
Causal base-rate usage (Stanovich & West, 1998c, 1998d)
Outcome bias (Stanovich & West, 1998c, 2008)
Framing within subjects
(Frederick, 2005; Parker & Fischhoff, 2005; Stanovich & West, 1998b, 1999)
Denominator neglect (Stanovich & West, 2008; Kokis et al., 2002)
Probability matching (Stanovich & West, 2008; West & Stanovich, 2003)
Hindsight bias (Stanovich & West, 1998c)
Ignoring P(D/NH) (Stanovich & West, 1998d, 1999)
Covariation detection (Stanovich & West, 1998c, 1998d; Sá et al., 1999)
Belief bias in syllogistic reasoning (Stanovich & West, 1998c, 2008)
Belief bias in modus ponens (Stanovich & West, 2008)
Informal argument evaluation (Stanovich & West, 1997, 2008)
Four-card selection task (Stanovich & West, 1998a, 2008)
EV maximization in gambles (Frederick, 2005; Benjamin & Shapiro, 2005)

irrational thought and behavior for its own ends (Blackmore, 1999; Distin, 2005; Stanovich, 2004, 2009). None of these factors are assessed on popular intelligence tests (or their proxies, like the SAT). Intelligence tests do not assess the *propensity* to override responses primed by the autonomous mind or to engage in full cognitive simulation. The crystallized abilities assessed on intelligence tests do not probe for the presence of the specific mindware that is critical for rational thought. And, finally, there are no probes on intelligence tests for the presence of contaminated mindware. Thus we should not be surprised when smart people act foolishly. That we in fact *are* sometimes surprised indicates that we are overvaluing and overconceptualizing the term *intelligence*—we are attributing to it qualities that intelligence tests do not measure. We are missing something important by treating intelligence as if it encompassed all cognitive abilities.

RATIONALITY CAN BE LEARNED AND IRRATIONALITY AMELIORATED

One of the things that we are missing is a focus on the malleability of rationality. This is ironic given that there are at least preliminary indications that rationality may be more malleable than intelligence.

Irrationality caused by mindware gaps is most easily remediable, as it is entirely due to missing strategies and declarative knowledge that can be taught. Overriding the tendencies of the autonomous mind is most often done with learned mindware, and sometimes override fails because of inadequately instantiated mindware. In such a case, inadequately learned mindware is the source of the problem. For example, disjunctive reasoning is the tendency to consider all possible states of the world in deciding among options or choosing a problem solution in a reasoning task. It is a rational thinking strategy with a high degree of generality. People make many suboptimal decisions because of the failure to flesh out all the possible options in a situation, yet the disjunctive mental tendency is not computationally expensive. This is consistent with the finding that there are not strong intelligence-related limitations on the ability to think disjunctively and with evidence indicating that disjunctive reasoning is a rational thinking strategy that can be taught (Adams, 1989; Baron & Brown, 1991; Feehrer & Adams, 1986; Kuhn, 2005; Nickerson, 1988, 2004; Reyna & Farley, 2006; Ritchhart & Perkins, 2005; Swartz & Perkins, 1989).

The tendency to consider alternative hypotheses is, like disjunctive reasoning, strategic mindware of great generality. Also, it can be implemented in very simple ways. Many studies have attempted to teach the importance of thinking of the alternative hypothesis by instructing people in a simple habit. People are given extensive practice at saying to themselves the phrase "think of the opposite" in relevant situations. This strategic mindware does not stress computational capacity and thus is probably easily learnable by many individuals. Several studies have shown that practice at the simple strategy of triggering the thought "think of the opposite" can help to prevent a host of the thinking errors studied in the literature of heuristics and biases, including but not limited to anchoring biases, overconfidence effects, hindsight bias, confirmation bias, and self-serving biases (Arkes, Fault, Guilmette, & Hart, 1988; Koehler, 1994; Koriat, Lichtenstein, & Fischhoff, 1980; Larrick, 2004; Mussweiler, Strack, & Pfeiffer, 2000).

Various aspects of probabilistic thinking represent mindware of great generality and potency. However, as any person who has ever taught a statistics course can attest (the present authors included), some of these insights are counterintuitive and unnatural for people—particularly in their application. There is nevertheless still some evidence that they are indeed teachable—albeit with somewhat more effort and difficulty than strategies such as disjunctive reasoning or considering alternative hypotheses. Aspects of scientific thinking necessary to infer a causal relationship are also definitely teachable (Kuhn, 2005, 2007; Leshowitz, DiCerbo, & Okun, 2002; Nisbett, 1993; Sedlmeier, 1999; Zimmerman, 2007).

Other strategies of great generality may be easier to learn—particularly by those of lower intelligence. For example, psychologist Peter Gollwitzer has discussed an action strategy of extremely wide generality—the use of implementation intentions (Gollwitzer, 1999; Gollwitzer & Schaal, 1998). An implementation intention is formed when the individual marks the cue–action sequence with the conscious, verbal declaration that "When X occurs, I will do Y." The triggering of this cue–action sequence on just a few occasions is enough to establish it in the autonomous mind. Finally, research has shown that an even more minimalist cognitive strategy of forming mental goals (whether or not they have implementation intentions) can be efficacious. For example, people perform better in a task when they are told to form a mental goal ("Set a specific, challenging goal for yourself") for their performance, as opposed to being given the generic motivational instructions ("Do your best"); (Heath, Larrick, & Wu, 1999; Locke & Latham, 1991).

Much of the strategic mindware discussed so far represents learnable strategies in the domain of instrumental rationality (achieving one's goals). Epistemic rationality (having beliefs well calibrated to the world) is often disrupted by contaminated mindware. However, even here, there are teachable macrostrategies that can reduce the probability of acquiring mindware harmful to its host. For example, the principle of falsifiability provides a wonderful inoculation against many kinds of nonfunctional beliefs. It is a tool of immense generality. It is taught in low-level courses on methodology and the philosophy of science, but could be taught much more broadly than this (Stanovich, 2010). Many pseudoscientific beliefs represent the presence of contaminated mindware. The critical thinking skills that help individuals to recognize pseudoscientific belief systems can be taught in high school courses.

If the cognitive miser is easily framed, responds to the most vivid stimulus present, and accepts defaults as given, then the behavior of misers will be shaped by whoever in their world has the power to determine these things (how things are framed, what the most vivid stimulus is, and what the default is). This is clearly problematic, but it suggests that there is another way (other than changing cognition directly) to help people avoid irrational acts. It suggests that a benevolent controller of our environment could help us—could save us from our irrational acts without our having to change basic aspects of our cognition. In short, for certain cognitive problems it might be easier to change the environment than to change people.

For example, in a cross-national study of organ donation rates, Johnson and Goldstein (2006) found that 85.9% of individuals in Sweden had agreed to be organ donors. However, the rate in the United Kingdom was only 17%. The difference in organ donorship between these countries has nothing to do with internal psychological differences between their citizens. The difference is due to a contrast in the public policy about becoming an organ donor in these different countries. In Sweden—like Belgium, France, Poland, and Hungary where agreement to organ donorship is over 95%—the default value on organ donorship is presumed consent. In countries with this public policy, people are assumed to have allowed their organs to be harvested but can opt out by taking an action (usually by getting a notation on their driver's licenses). In contrast, the United States and United Kingdom—like Germany, Denmark, and the Netherlands where agreement to organ donorship is less than 30%—the default value is no donation, with explicit action required to opt *for* organ donation.

In short, the difference between Sweden and the United Kingdom is not in the people. The citizens of both countries are cognitive misers and probably to a roughly equal extent. The great difference is in the form of a particular public policy. As misers, the citizens of both countries are strongly affected by the default heuristic. The option offered as the default is “sticky” in that it is overly influential. A very small change in the donor decision-making environment that hurts no one (since an opt-out procedure is allowed in all countries with presumed consent) could save the lives of thousands of people. The tendencies of the cognitive miser have cost thousands of people their lives. But these tragic consequences are preventable. The best prevention in this case, though, is a change in the environment rather than a change in people, because the former is so much easier to implement.

Thaler and Sunstein (2008) provide another example of environmental fixes by discussing pension participation reforms. The first step comes at the point when employees of most large companies must first choose to enroll. If they do nothing (do not fill out the relevant form) they are not enrolled. Here is where things first go wrong. Many employees do not enroll. New reforms have the employees automatically signed up for the 401(k). They must choose (by filling out a form) to opt out of the system. Such a reform exploits the default bias of the cognitive miser.

An even larger category of problems where people need help from their environments involves their self-control. People overeat, overspend, procrastinate, smoke, and drink too much. Solutions to these problems with self-control are of two forms, corresponding to changes in the individual and changes in the environment. People try to bolster their “will power”—that is, their internal powers of self-control. Alternatively, they try to rearrange their environments so that less exercise of will power (autonomous system override) will be necessary. A common strategy here is to use precommitment devices. People enroll in automatic savings plans so that they will not overspend. They prepackage meals so that they will not overeat. They commit themselves to deadlines so that they will not procrastinate. Precommitments represent our deliberate attempts to restructure our environments so that they will be more conducive to our attempts at self-control.

All of these examples show how simple environmental changes can prevent problems in rational thinking problems. Many more such examples are discussed in Gigerenzer (2002), Stanovich (2009), and Thaler and Sunstein (2008).

RATIONALITY ENCOMPASSES CRITICAL THINKING AND INTELLIGENCE

We can tame intelligence in folk psychology by pointing out that there are legitimate scientific terms as well as folk terms for the other valued parts of cognitive life and that some of these are measurable. This strategy uses to advantage a fact of life that many critics of IQ tests have lamented—that intelligence tests are not going to change any time soon. The tests have the label "intelligence"; thus what they measure will always be dominant in the folk psychology of intelligence. I would argue that it is a mistake to ignore this fact. The tests do not measure rationality, and thus the ability to think rationality will be a subordinate consideration in our schools, in our employment selection devices, and in our society as a whole.

Stanovich (2009) has argued for opening up some space for rationality in the lexicon of the mental and, in doing so, tame the intelligence concept. The term dysrationalia (an analogue of the word *dyslexia*) was defined as the inability to think and behave rationally despite having adequate intelligence (see Stanovich, 1993). Of course, it is easy to recognize that this definition was formulated to contain linguistic and conceptual parallels with the learning disability definitions that stress aptitude–achievement discrepancy. The idea of defining a disability as an aptitude–achievement discrepancy (performance on some domain that is unexpectedly below intelligence) spread widely during the early years of the development of the learning disability concept. Note that the discrepancy idea contains the assumption that all good things should go with high intelligence. When a high score on an IQ test is accompanied by subpar performance in some other domain, this is thought "surprising," and a new disability category is coined to name the surprise.

The strategy in proposing dysrationalia was to prevent intelligence from absorbing the concept of rationality—something that IQ tests do not measure. Restricting the term *intelligence* to what the tests actually measure has the advantage of getting usage in line with the real world of measurement and testing. We have coherent and well-operationalized concepts of rational action and belief formation. We have a coherent and well-operationalized concept of intelligence. No scientific purpose is served by fusing these concepts, because they are very different. To the contrary, scientific progress is made by *differentiating* concepts. Dysrationalia highlights the fact that "all good things" (rationality in this case) do not always go with intelligence. The concept of dysrationalia (and the empirical evidence indicating that the condition is not rare) should help to attenuate our surprise at intelligence–rationality dissociations and to create conceptual space in which we can value abilities at least as important as those presently measured on IQ tests—abilities to form rational beliefs and to take rational action.

The tripartite model of mind presented in this chapter explains why rationality is a more encompassing construct than intelligence. Rationality requires the proper functioning of both the reflective and the algorithmic mind. In contrast, intelligence tests index the computational power of the algorithmic mind. Likewise, the construct of critical thinking is subsumed under the construct of rationality. For example, the processes of critical thinking are often summarized as a set of thinking dispositions that must be developed or inhibited: need for cognition, actively open-minded thinking, belief identification, consideration of future consequences, reflectivity/impulsivity, rational/experiential orientation, need for closure, openness, conscientiousness, etc. These thinking dispositions are the individual difference constructs that capture the functioning of the reflective mind in the tripartite model.

In the context of this model, rationality requires three things: the propensity to override suboptimal responses from the autonomous mind, the algorithmic capacity to inhibit the suboptimal response and to simulate an alternative, and the presence of the mindware that allows the computation of an alternative response. The propensity to override suboptimal responses from the autonomous mind—a property of the reflective mind—captures virtually all of the propensities of critical thinking that have been discussed in the traditional literature on that construct. The algorithmic capacity to inhibit the suboptimal response and to simulate an alternative is captured in standard intelligence tests. Thus, folding critical thinking and intelligence into a generic model of the mind that has rationality as an overarching construct has the considerable advantage of situating, virtually for the first time, the construct of critical thinking within contemporary cognitive science.

REFERENCES

Ackerman, P. L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. 4, pp. 1–27). Norwood, NJ: Ablex.

Ackerman, P. L. (1996). A theory of adult development: Process, personality, interests, and knowledge. *Intelligence, 22,* 227–257.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121,* 219–245.

Ackerman, P. L., & Kanfer, R. (2004). Cognitive, affective, and conative aspects of adult intellect within a typical and maximal performance framework. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 119–141). Mahwah, NJ: Erlbaum.

Adams, M. J. (1989). Thinking skills curricula: Their promise and progress. *Educational Psychologist, 24,* 25–77.

Adams, M. J. (1993). Towards making it happen. *Applied Psychology: An International Review, 42,* 214–218.

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Anderson, M. (2005). Marrying intelligence and cognition: A developmental view. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence* (pp. 268–287). New York: Cambridge University Press.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110,* 486–498.

Arkes, H., Faust, D., Guilmette, T., & Hart, K. (1988). Eliminating the hindsight bias. *Journal of Applied Psychology, 73,* 305–307.

Audi, R. (1993). *The structure of justification.* Cambridge, UK: Cambridge University Press.

Audi, R. (2001). *The architecture of reason: The structure and substance of rationality.* Oxford, UK: Oxford University Press.

Austin, E. J., & Deary, I. J. (2002). Personality dispositions. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 187–211). New Haven, CT: Yale University Press.

Baron, J. (1982). Personality and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 308–351). Cambridge, UK: Cambridge University Press.

Baron, J. (1993). Why teach thinking?—An essay. *Applied Psychology: An International Review, 42,* 191–214.

Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences, 17,* 1–42.

Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning, 1,* 221–235.

Baron, J. (2008). *Thinking and deciding* (4th ed.). New York: Cambridge University Press.

Baron, J., & Brown, R. V. (Eds.). (1991). *Teaching decision making to adolescents.* Hillsdale, NJ: Erlbaum.

Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: The relationship of inspection time and openness to g and Gc. *Intelligence, 31,* 275–287.

Bermudez, J. L. (2001). Normativity and rationality in delusional psychiatric disorders. *Mind & Language, 16,* 457–493.

Blackmore, S. (1999). *The meme machine.* New York: Oxford University Press.

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (Vol. 28, pp. 41–100). San Diego, CA: Academic Press.

Bruine de Bruin, W., Parker, A.M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology, 92,* 938–956.

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119,* 197–253.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1–22.

Cattell, R. B. (1998). Where is intelligence? Some answers from the triadic theory. In J. J. McArdle & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 29–38). Mahwah, NJ: Erlbaum.

Cronbach, L. J. (1949). *Essentials of psychological testing.* New York: Harper.

Cuban, L. (1984). Policy and research dilemmas in the teaching of reasoning: Unplanned designs. *Review of Educational Research, 54,* 655–681.

Dawes, R. M. (1976). Shallow psychology. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 3–11). Hillsdale, NJ: Erlbaum.

Dawes, R. M. (1998). Behavioral decision making and judgment. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 1, pp. 497–548). New York: McGraw-Hill.

De Neys, W. (2006). Dual processing in reasoning—Two systems but one reasoner. *Psychological Science, 17,* 428–433.

Dennett, D.C. (1996). *Kinds of minds: Toward an understanding of consciousness.* New York: Basic Books.

Distin, K. (2005). *The selfish meme.* Cambridge, UK: Cambridge University Press.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist, 33,* 109–128.

Donaldson, M. (1978). *Children's minds.* London: Fontana Paperbacks.

Donaldson, M. (1993). *Human minds: An exploration.* New York: Viking Penguin.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51,* 380–417.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron, & R. Sternberg (Eds.). *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: Freeman.

Ennis, R. H. (1996). Critical thinking dispositions: Their nature and assessability. *Informal Logic, 18,* 165–182.

Evans, J. St. B. T. (1972). Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology, 24,* 193–199.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences.* Hove, UK: Erlbaum.

Evans, J. St. B. T. (1998). Matching bias in conditional reasoning: Do we understand it after 25 years? *Thinking and Reasoning, 4,* 45–82.

Evans, J. St. B. T. (2002). The influence of prior belief on scientific thinking. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 193–210). Cambridge, UK: Cambridge University Press.

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454–459.

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review, 13,* 378–395.

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment.* New York: Psychology Press.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology, 59,* 255–278.

Evans, J. St. B. T. (2009). How many dual-process theories do we need? One, two, or many? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33–54). Oxford, UK: Oxford University Press.

Evans, J. St. B. T., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11,* 295–306.

Evans, J. St. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking and Reasoning, 11,* 382–389.

Evans, J. St. B. T., & Feeney, A. (2004). The role of prior belief in reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 78–102). Cambridge, UK: Cambridge University Press.

Evans, J. St. B. T., & Lynch, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology, 64,* 391–397.

Evans, J. St. B. T., Newstead, S., Allen, J., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology, 6,* 263–285.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove, UK: Psychology Press.

Evans, J. St. B. T., & Over, D. E. (2004). *If.* Oxford, UK: Oxford University Press.

Eysenck, H. J. (1994). Personality and intelligence: Psychometric and experimental approaches. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 3–31). Cambridge, UK: Cambridge University Press.

Feehrer, C. E., & Adams, M. J. (1986). *Odyssey: A curriculum for thinking.* Watertown, MA: Charlesbridge.

Feldman Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130,* 553–573.

Foley, R. (1987). *The theory of epistemic rationality.* Cambridge, MA: Harvard University Press.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19,* 25–42.

Garnham, A., & Oakhill, J. V. (2005). Accounting for belief bias in a mental model framework: Comment on Klauer, Musch, and Naumer (2000). *Psychological Review, 112,* 509–518.

Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you.* New York: Simon & Schuster.

Gilhooly, K. J., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking and Reasoning, 11,* 279–302.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology, 84,* 537–552.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54,* 493–503.

Gollwitzer, P. M., & Schaal, B. (1998). Metacognition in action: The importance of implementation intentions. *Personality and Social Psychology Review, 2,* 124–136.

Greene, J., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44,* 389–400.

Hardman, D. (1998). Does reasoning occur on the selection task? A comparison of relevance-based theories. *Thinking and Reasoning, 4,* 353–376.

Harman, G. (1995). Rationality. In E. E. Smith & D. N. Osherson (Eds.), *Thinking* (Vol. 3, pp. 175–211). Cambridge, MA: MIT Press.

Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York: Oxford University Press.

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world.* Thousand Oaks, CA: Sage.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38,* 79–109.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26,* 1–23.

Jeffrey, R. C. (1983). *The logic of decision* (2nd ed.). Chicago: University of Chicago Press.

Johnson, E. J., & Goldstein, D. G. (2006). Do defaults save lives? In S. Lichtenstein & P. Slovic (Eds.), *The construction of preference* (pp. 682–688). Cambridge, UK: Cambridge University Press.

Johnson-Laird, P. N. (2006). *How we reason.* Oxford, UK: Oxford University Press.

Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes. *Behavioral and Brain Sciences, 23,* 681–683.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697–720.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267–293). New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames.* Cambridge, UK: Cambridge University Press.

Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review, 111,* 512–523.

Kelley, D. (1990). *The art of reasoning.* New York: Norton.

Klaczynski, P. A. (1997). Bias in adolescents' everyday reasoning and its relationship with intellectual ability, personal theories, and self-serving motivation. *Developmental Psychology, 33,* 273–283.

Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology, 89,* 470–485.

Klaczynski, P. A., & Lavallee, K. L. (2005). Domain-specific identity, epistemic regulation, and intellectual ability as predictors of belief-based reasoning: A dual-process perspective. *Journal of Experimental Child Psychology, 92,* 1–24.

Klaczynski, P. A., & Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: Adult age differences in everyday reasoning tasks. *Psychology and Aging, 15,* 400–416.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*, 852–884.

Koehler, D. J. (1994). Hypothesis generation and confidence in judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 461–469.

Kokis, J., Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83,* 26–52.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing the mind: "Seizing" and "freezing." *Psychological Review, 103,* 263–283.

Kuhn, D. (1993). Connecting scientific and informal reasoning. *Merrill-Palmer Quarterly, 38,* 74–103.

Kuhn, D. (2001). How do people know? *Psychological Science, 12,* 1–8.

Kuhn, D. (2005). *Education for thinking.* Cambridge, MA: Harvard University Press.

Kuhn, D. (2007, Feb./March). Jumping to conclusions: Can people be counted on to make sound judgments? *Scientific American Mind,* 44–51.

Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning, 13,* 90–104.

Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment & decision making* (pp. 316–337). Malden, MA: Blackwell.

LeBoeuf, R. A., & Shafir, E. (2005). Decision making. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 243–265). New York: Cambridge University Press.

Leshowitz, B., DiCerbo, K. E., & Okun, M. A. (2002). Effects of instruction in methodological reasoning on information evaluation. *Teaching of Psychology, 29,* 5–10.

Leslie, A. M. (1987). Pretense and representation: The origins of "Theory of Mind." *Psychological Review, 94,* 412–426.

Levesque, H. J. (1986). Making believers out of computers. *Artificial Intelligence, 30,* 81–108.

Levesque, H. J. (1989). Logic and the complexity of reasoning. In R. H. Thomason (Ed.), *Philosophical logic and artificial intelligence* (pp. 73–107). Kluwer.

Lieberman, M. D. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes* (pp. 44–67). New York: Cambridge University Press.

Lipman, M. (1991). *Thinking in education.* Cambridge, UK: Cambridge University Press.

Locke, E. A., & Latham, G. P. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes, 50,* 212–247.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions.* New York: Wiley.

Macpherson, R., & Stanovich, K. E. (2007). Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences, 17,* 115–127.

Manktelow, K. I. (2004). Reasoning and rationality: The pure and the practical. In K. I. Manktelow & M. C. Chung (Eds.), *Psychology of reasoning: Theoretical and historical perspectives* (pp. 157–177). Hove, UK: Psychology Press.

Margolis, H. (1987). *Patterns, thinking, and cognition.* Chicago: University of Chicago Press.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science & myth.* Cambridge, MA: MIT Press.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306,* 503–507.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of will power. *Psychological Review, 106,* 3–19.

Miyake, A., Friedman, N., Emerson, M. J., & Witzki, A. H. (2000). The utility and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

Moshman, D. (2004). From inference to reasoning: The construction of rationality. *Thinking and Reasoning, 10,* 221–239.

Moshman, D. (2005). *Adolescent psychological development: Rationality, morality, and identity* (2nd ed.). Mahwah, NJ: Erlbaum.

Moshman, D. (2010). The development of rationality. In H. Siegel (Ed.), *Oxford Handbook of Philosophy of Education* (pp. 145–161). Oxford, UK: Oxford University Press.

Murphy, D., & Stich, S. (2000). Darwin in the madhouse: Evolutionary psychology and the classification of mental disorders. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the human mind: Modularity, language and meta-cognition* (pp. 62–92). Cambridge, UK: Cambridge University Press.

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin, 9,* 1142–1150.

Neimark, E. (1987). *Adventures in thinking.* San Diego, CA: Harcourt Brace Jovanovich.

Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology, 57A,* 33–60.

Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds.* Oxford, UK: Oxford University Press.

Nickerson, R. (1987). Why teach thinking? In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 27–40). New York: Freeman.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175–220.

Nickerson, R. S. (2004). *Cognition and chance: The psychology of probabilistic reasoning.* Mahwah, NJ: Erlbaum.

Nisbett, R. E. (1993). *Rules for reasoning.* Hillsdale, NJ: Erlbaum.

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Midwest Publications.

Over, D. E. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 3–18). Malden, MA: Blackwell Publishing.

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual differences approach. *Journal of Behavioral Decision Making, 18,* 1–27.

Paul, R. W. (1984). Critical thinking: Fundamental to education for a free society in North America: A new theory of knowledge, learning, and literacy. *Educational Leadership, 42,* 4–14.

Paul, R. W. (1987). Critical thinking and the critical person. In D. N. Perkins, J. Lockhead & J. Bishop (Eds.), *Thinking: The second international conference* (pp. 373–403). Hillsdale, NJ: Erlbaum.

Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence.* New York: Free Press.

Perkins, D. N. (2002). The engine of folly. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 64–85). New Haven, CT: Yale University Press.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Piaget, J. (1926). *The language and thought of the child.* London: Routledge.

Reyna, V. F. (2004). How people make decisions that involve risk. *Current Directions in Psychological Science, 13,* 60–66.

Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making. *Psychological Science in the Public Interest, 7,* 1–44.

Ritchhart, R., & Perkins, D. N. (2005). Learning to think: The challenges of teaching thinking. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 775–802). New York: Cambridge University Press.

Sá, W., Kelley, C., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: Individual differences in the coordination of theory and evidence. *Personality and Individual Differences, 38,* 1149–1161.

Sá, W., & Stanovich, K. E. (2001). The domain specificity and generality of mental contamination: Accuracy and projection in judgments of mental content. *British Journal of Psychology, 92,* 281–302.

Sá, W., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91,* 497–510.

Samuels, R., & Stich, S. P. (2004). Rationality and psychology. In A. R. Mele & P. Rawling (Eds.), *The Oxford handbook of rationality* (pp. 279–300). Oxford, UK: Oxford University Press.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist, 39,* 19–30.

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications.* Mahwah, NJ: Erlbaum.

Shafir, E. (1994). Uncertainty and the difficulty of thinking through disjunctions. *Cognition, 50,* 403–430.

Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology, 53,* 491–517.

Siegel, H. (1988). *Educating reason.* New York: Routledge.

Siegel, H. (1997). *Rationality redeemed? Further dialogues on an educational ideal.* New York: Routledge.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22.

Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 379–396). New York: Cambridge University Press.

Stanovich, K. E. (1993). Dysrationalia: A new specific learning disability. *Journal of Learning Disabilities, 26,* 501–515.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning.* Mahwah, NJ: Erlbaum.

Stanovich, K. E. (2002). Rationality, intelligence, and levels of analysis in cognitive science: Is dysrationalia possible? In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 124–158). New Haven, CT: Yale University Press.

Stanovich, K. E. (2003). The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 291–342). New York: Cambridge University Press.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin.* Chicago: University of Chicago Press.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought.* New Haven, CT: Yale University Press.

Stanovich, K. E. (2010). *How to think straight about psychology* (9th ed.). Boston: Allyn & Bacon.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89,* 342–357.

Stanovich, K. E., & West, R. F. (1998a). Cognitive ability and variation in selection task performance. *Thinking and Reasoning, 4,* 193–230.

Stanovich, K. E., & West, R. F. (1998b). Individual differences in framing and conjunction effects. *Thinking and Reasoning, 4,* 289–317.

Stanovich, K. E., & West, R. F. (1998c). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127,* 161–188.

Stanovich, K. E., & West, R. F. (1998d). Who uses base rates and P(D/~H)? An analysis of individual differences. *Memory & Cognition, 26,* 161–179.

Stanovich, K. E., & West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. *Cognitive Psychology, 38,* 349–385.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645–726.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning, 13,* 225–247.

Stanovich, K. E., & West, R. F. (2008a). On the failure of intelligence to predict myside bias and one-sided bias. *Thinking & Reasoning, 14,* 129–167.

Stanovich, K. E., & West, R. F. (2008b). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672–695.

Sternberg, R. J. (1997). *Thinking styles.* Cambridge, UK: Cambridge University Press.

Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist, 36,* 227–245.

Sternberg, R. J. (2002). Smart people are not stupid, but they sure can be foolish: The imbalance theory of foolishness. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 232–242). New Haven, CT: Yale University Press.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized.* Cambridge, UK: Cambridge University Press.

Sternberg, R. J. (2005). Foolishness. In R. J. Sternberg & J. Jordan (Eds.), *A handbook of wisdom: Psychological perspectives* (pp. 331–352). New York: Cambridge University Press.

Sternberg, R. J. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34,* 321–350.

Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style? *American Psychologist, 52,* 700–712.

Sternberg, R. J., & Jordan, J. (Eds.). (2005). *A handbook of wisdom: Psychological perspectives.* New York: Cambridge University Press.

Strathman, A., Gleicher, F., Boninger, D. S., & Scott Edwards, C. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66,* 742–752.

Swartz, R. J., & Perkins, D. N. (1989). *Teaching thinking: Issues & approaches.* Pacific Grove, CA: Midwest Publications.

Taylor, S. E. (1981). The interface of cognitive and social psychology. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 189–211). Hillsdale, NJ: Erlbaum.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Toplak, M., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K. E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making, 20,* 103–124.

Toplak, M. E., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 94,* 197–209.

Toplak, M. E., & Stanovich, K. E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology, 17,* 851–860.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211,* 453–458.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59,* 251–278.

Tweney, R. D., & Yachanin, S. (1985). Can scientists rationally assess conditional inferences? *Social Studies of Science, 15,* 155–173.

von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Voss, J., Perkins, D., & Segal J. (Eds.). (1991). *Informal reasoning and education.* Hillsdale, NJ: Erlbaum.

Wade, C., & Tavris, C. (1993). *Critical and creative thinking.* New York: HarperCollins.

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, UK: Penguin.

West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition, 31,* 243–251.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117–142.

Wu, G., Zhang, J., & Gonzalez, R. (2004). Decision under risk. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 399–423). Malden, MA: Blackwell.

Zechmeister, E. B., & Johnson, J. (1992). *Critical thinking: A functional approach.* Pacific Grove, CA: Brooks/Cole.

Zeidner, M., & Matthews, G. (2000). Intelligence and personality. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 581–610). New York: Cambridge University Press.

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8,* 12–17.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27,* 172–223.

This page intentionally left blank

PART IV

Cultural Aspects of Teaching

This page intentionally left blank

9

A Theory of Teaching

ALAN H. SCHOENFELD

There are at least two major reasons for studying teaching. The first is that teaching is a highly complex, socially interactive activity. If we, as a field, can understand how and why teachers make the choices they make while engaged in classroom instruction, we will have come a long way in understanding the mechanisms that underlie complex human decision making. The second is that the more deeply one understands the teaching process, the more one may be able to improve it.

I entered educational/psychological research more than 30 years ago as a mathematician with a deep love of mathematics and a wish to find ways to help students experience the pleasures of mathematical thinking as I knew them. My goal throughout my career has been to develop deeper understandings of thinking, teaching, and learning in order to make the pleasures of thinking mathematically more accessible to as many people as possible while at the same time constructing increasingly deep models of cognitive processes. This chapter provides a brief overview of the evolution of my ideas, the current state of the theory of teaching-in-context, and the current and prospective practical applications of that theory.

NOT LONG AGO, A VERY DIFFERENT WORLD

In many ways, advances in our understanding of thinking and learning over the past half-century (even the past 25 years!) have been absolutely breathtaking—so much so that the intellectual landscape looks radically different now, just as the technological landscape does. For example, in the early 1970s the personal computer, much less the web, did not exist. My PhD dissertation was produced on a typewriter in 1973; the Apple I computer, with a grand total of 4KB of RAM, was introduced in 1976. Today, when anyone with a cell phone can be reached just about anywhere, any time, via voice or text messaging, it is difficult for those who have grown up living with such technologies to imagine a world without them. Similarly, it is difficult for today's students to imagine a time when metacognition and belief systems did not exist—as constructs, that is. Thus, as context for what follows, I begin by characterizing the state of the art when I began doing educational research.

Mathematics education was just beginning to coalesce as a discipline in the late 1960s and early 1970s. The emergence of the discipline represents an important development, signaling recognition of the importance of domain-specific content knowledge and (mathematical) epistemology above and beyond general theories of learning. The journal *Educational Studies in Mathematics,* edited by Hans Freudenthal, made its debut in Holland in 1968; the first International Congress on Mathematics Education took place in Lyon, France, in 1969; the German journal *Zentralblatt für Didaktik der Mathematik (ZDM)* appeared in 1969; and in the United States, the *Journal for Research in Mathematics Education* published its first issue in 1970. Thus, despite its distinguished lineage in philosophy and psychology, one can argue that mathematics education as a field is only 40 years old.

Not only that, but in the 1960s and 1970s the broader field (including psychology) was decidedly noncognitive in that it paid little heed to cognitive processes as active objects of investigation. Americans were aware of Piaget's work but usually in a reductive way. Piaget's developmental stages were typically taken as *constraints:* for example the Piagetian claim that children do not become formal reasoners until they reach their teenage years was used by some scholars as an argument that students in the elementary grades should not be exposed to tasks that called for formal reasoning. Constructs that we now take for granted—such as metacognition and belief systems—did not yet exist. Educational/psychological research methods in the United States were largely statistical/experimental and modeled simplistically on research in areas such as agriculture. If two adjacent and otherwise identical fields of corn were treated identically except for one variable (say the amount of watering or the use of two different fertilizers), different yields of corn could be reasonably attributed to the differences in treatment. Similarly, if two randomly selected (and thus presumably equivalent) groups of students were given two different instructional "treatments," the differential performance of the two groups could presumably be attributed to the two different treatments. (This is a gross oversimplification: see Schoenfeld, 2000, 2007, for detail.) Psychology was, in a sense, still emerging from the throes of behaviorism, according to which the very notion of cognitive processes ("mentalism") was a theoretical heresy.

The seeds of change came from many directions, most notably the set of disciplines that coalesced to form cognitive science—among them artificial intelligence, psychology, philosophy, linguistics, and education. For example, a core aspect of artificial intelligence (AI) is that close observations of people engaged in problem solving can reveal systematicity in successful problem-solving attempts—and that this can be captured in computer programs (see, e.g., Newell & Simon, 1972). Although work in AI was aimed at producing *machine intelligence*—computer programs that did problem solving and such, it had the ironic consequence of, at least theoretically, putting behaviorism to rest. The behaviorists, who claimed that the only phenomena that counted were observable outcomes, were hoist by their own theoretical petard when AI programs, derived on the basis of "mentalist" theories of problem solving, turned out to be effective problem solvers! AI provided a form of validation for information processing psychology and also a metaphor for the kind of work I wound up doing: if we could learn a lot about effective *machine* problem solving from careful observations of humans solving problems, it stood to reason that we could learn a lot about effective *human* problem solving from careful observations of humans solving problems. That is where I got started—with an attempt to describe effective problem-solving strategies for human rather than computer implementation.

The idea was simple. George Pólya, one of the 20th century's great mathematicians, had written extensively about problem-solving strategies, which he called *heuristics*—roughly speaking, rules of thumb for making progress on or developing a better understanding of challenging mathematical problems (Pólya, 1945, 1954, 1962, 1965). As a practicing mathematician, I had read Pólya's descriptions of problems and resonated to them; so had other mathematicians and mathematics educators.

But attempts to teach problem solving "à la Pólya" had not been successful. That the descriptions of strategies felt right but did not seem to help students led me to ask: Might one use the techniques of AI to clarify the strategies discussed by Pólya, so that mathematics learners (rather than computers, as in AI) could learn to use the strategies? That is, humans and computers are both information processors, but with very different capacities. In AI, one looked for strategies that could be implemented by computers, which have extraordinarily large and essentially infallible memory and spectacular (and essentially error-free) computational capacity but poor linguistic capacity and pattern recognition. In contrast, humans have good but not great long-term memories, severe limitations in short-term memory (the "magic number seven plus or minus two"; see Miller, 1956), and computational capacity that, on average, is mediocre at best. But humans compensate for those limitations with remarkable linguistic and pattern recognition abilities. What if one observed people solving problems with an eye toward abstracting strategies that other people could use?

In short, this approach bore fruit; and as it did, the field of cognitive science coalesced as a discipline. By 1985 researchers in a number of fields had codified the dimensions of problem-solving performance in mathematics and other goal-oriented analytic disciplines (e.g., writing). In the mid-1970s I began by explicating problem-solving strategies at a finer level of detail than had been examined previously, adapting the ideas of Newell and Simon. Heuristic strategies turned out to be learnable but not directly à la Pólya: ostensibly simple strategies such as "exploit easier related problems" turned out to be families of related strategies, all of which needed to be learned before one could be said to have learned the "simple" strategy. This led to studies of metacognition—how does one "manage" myriad strategies? This, in turn, led to studies of belief systems—why do people sometimes make what seem to be rather strange choices when they are engaged in problem solving? The core result, after a decade of research and development (I taught courses in problem solving and analyzed students' behavior) was a framework for the analysis of problem-solving behavior. Details can be found in my 1985 book *Mathematical Problem Solving*.

Mathematical Problem Solving summed things up as follows. In order to understand the reasons for the success or failure of people's attempts at mathematical problem solving, one must examine their knowledge base; access to problem-solving strategies; metacognition, specifically monitoring and self-regulation; and beliefs and practices.

Those four categories are necessary, in that failure to take any of them into account in examining a problem-solving attempt can result in missing the cause of success or failure. They are sufficient for analysis in that no additional categories of analysis are necessary.

This framework, which can be used in a wide range of disciplines to explain an individual's success or failure to solve a given problem, has stood the test of time. In that sense it is robust; but it is also limited. Use of the framework does serve to explain why a problem-solving attempt succeeds or fails. However, it does not help to explain how and why the problem solver made particular decisions along the way—why *this* knowledge rather than *that* knowledge was accessed and used, why a particular solution path was adopted or avoided, and so on. In short, what was lacking was a theory of decision making at the level of action—a theory of how and why the problem solver, in the moment, chooses to do what he or she does.

STUDIES OF TUTORING AS A BRIDGE TO STUDIES OF TEACHING

In studies of individual problem solving, certain things are simplified—notably social interactions and goal structure. Social interactions are still present in individual problem solving. For example, there are the "demand characteristics" of tasks, which can affect people's behavior very strongly (see chapter 10 of *Mathematical Problem Solving*). However, the behavior of an individual acting alone is markedly different (and less complex) than the behavior of a person interacting with others. Moreover, in interactions goals can change rapidly. In problem solving, the main goal—solving the problem—typically remains unchanged and drives much of the individual's behavior.

In a move toward explaining phenomena of increased complexity, my research group turned to studies of tutoring. Tutoring is a knowledge-intensive activity with the specific goal (problem) of getting someone else to understand a particular body of content. Thus tutoring can be seen as a complex form of problem solving. But unlike the case in problem solving, the problem state during tutoring can change rapidly and often, as when a student reveals a previously unsuspected misconception ("What now? Do I remediate, ignore, or . . . ?") or gets the "deer in the headlights" look when asked what seemed to be a simple question. Although tutoring sessions are nowhere near as complex interactionally or cognitively as whole-class teaching, the study of tutoring represents a step up in complexity and a step toward understanding decision making during teaching.

At the time we began our studies, there were two separate and very different literatures on tutoring. One, grounded largely in artificial intelligence and intelligent tutoring systems, focused on building models of student cognition and figuring out the best "next step" to move the student toward deeper understanding. By and large, these studies paid little to no attention to issues of student affect during tutoring sessions. A second focused on "human factors," including motivation. By and large, these studies paid little or no attention to issues of the content being discussed during tutoring sessions. That had to be wrong. Any tutor knows that both affect and content can be highly consequential during a tutoring session and can shape what the tutor does. Things can be going along swimmingly when a misconception arises or the student suddenly indicates frustration or fatigue; when that happens, the tutor may have to change direction right then and there. Thus, what we sought was an "architecture" of decision making that unified content and affective concerns and did not automatically privilege one over the other.[1]

It turned out that the key to unification was what we called a "goal-driven architecture." When something happens—be it content- or affect-related—the tutor must decide whether it is important enough to deal with, and if so, how. If the answer is no, then things go on as planned, though the plans for later on may be changed. If the answer is yes, then there is a new goal in place: deal with this situation. How the tutor deals with the situation depends on what the tutor believes is important (e.g., is it OK to correct a mistaken procedure by saying, say, "careful, don't forget the middle terms when you calculate $(a + b)^2$. It's $a^2 + 2ab + b^2$, not $a^2 + b^2$," or is it important to make sure that the student really understands the multiplication?); it also depends on what knowledge is accessible to the tutor. The tutor may choose to "tell" or "delve" depending on how important he or she feels it is to root out the problem, how much time there is, and whether there is a readily accessible way to do it. In any case the situation will be dealt with by the tutor's drawing on his or her knowledge base. Once the situation has been dealt with, the tutor picks the next major goal (typically, when there is an interruption, a return to the original agenda) and moves forward.

The sketch above outlines the key elements in a theory of tutoring (which, I have to admit, became clearer with hindsight, after working on teaching for some years). They are highlighted below in italics.

The tutor enters a tutoring session with an *agenda,* which may be as simple as to work through a set of problems or may be more complex.[2] Explicitly or tacitly, the agenda defines a set of *goals*—to work through the first problem (or to work on the first major goal) until the tutor is satisfied that the issues it raises have been satisfactorily dealt with, at which point that goal is declared satisfied and work begins on the next major goal.

The tutor enters the session with a body of *knowledge* (about mathematics, about the student, etc.) that serves as the reservoir from which particulars will be drawn for interactions with the student. As in most everything else humans do, this knowledge is organized in the form of schemata, scripts, routines, and so forth. Thus, if the goal is to work through a related rates problem (or fraction addition, or any other problem), the tutor has ways of working these, ways of talking about what is important in the problem, and ways of finding out what the student knows and interacting with the student. These get activated when the goal of working through a particular problem is activated, and the tutor must choose among them.

The tutor enters the session with a body of *beliefs* (about mathematics, about the student, about what constitute appropriate tutorial/interpersonal interactions with the tutee, etc.) that shape the tutor's decision making. For example, a tutor who believes that students should have clear models to follow may give a "minilecture" or lay out a clean solution to a problem for a student at the beginning of an interaction, while a tutor who believes that interactions should begin with informal diagnoses of what the student does and does not understand is likely to begin differently. Similarly, beliefs about what is important in the mathematics will shape what the tutor chooses to emphasize.

Finally, there is *monitoring and decision making.* The default is that things will go as organized by the agenda: In a mathematics tutoring session, one may start with small talk, an overview of the topic/agenda, and then start working on a particular problem. The way the tutor and student work through the problem depends on the tutor's beliefs and knowledge, of course. Typically, tutors have scripts or routines for working through problems and the problems have natural subgoal structures, so there is typically a straightforward way to work through a problem. When a problem or a topic is done, that occasions a move to the next problem or topic. Of course, things do not always go according to plan. At any time, the student may say or do something that merits attention—perhaps requiring a minor comment or adjustment, perhaps (as in the case when a major content or affective issue arises) requiring a significant change of goals and plans. Thus *emergent goals* enter the goal stack, at which point things return to "normal" in terms of the architecture: tutor and student work toward the satisfaction of the top-level goal, and when it is satisfied, they return to the next goal on the stack (quite possibly the goal that was interrupted by the issue that halted the flow of events).

This goal-driven architecture, with adequate attention given to the tutor's knowledge and beliefs, yielded rich descriptions of tutors' decision making, and managed to unify the two literatures (intelligent tutoring systems and human factors) about tutoring (Arcavi & Schoenfeld, 1992; Schoenfeld et al., 1992). But it is a large leap from one-on-one tutoring sessions to the chaos of a classroom, where the teacher typically has multiple agendas and dozens of students to attend to. The question is whether a similar kind of approach could explain teachers' in-the-moment decision making, and if so, what the use of such a theory would be.

A THEORY OF TEACHING-IN-CONTEXT, PART 1: EXAMPLES OF KEY CONSTRUCTS

In broadest terms, my research on teaching has been aimed at the following:

> Imagine that you have had access to a teacher for some time, so that you understand what the teacher values and believes is important, what the teacher knows (in terms of content and pedagogy and about his or her students), and what he or she is trying to achieve in the classroom. Suppose now that you are watching a lesson unfold. Would it be possible to explain the teacher's actions in-the-moment on the basis of what you know? Might one even be able to predict, in a disciplined rather than ad hoc way, the way the teacher will react (and why)?

Here are three cases in point. In each I explain a situation where a teacher confronts or makes a decision. In what follows I will explain the teachers' decision making.

Case 1

MN, a beginning teacher, has orchestrated a classroom discussion of reducing algebraic functions. He has had the class work individually on reducing the fractions in $\frac{m^6}{m^2}, \frac{x^3y^7}{x^2y^6}$, and $\frac{x^5}{x^5}$. He walks around the classroom, monitoring student work on the problems, and then convenes the class for discussion. As he expects, the discussion of the first two problems goes smoothly. He calls on students who have found the right answers and asks for explanations; then he interactively elaborates on their explanations for the whole class. He expects students to have some difficulty with the third problem, and they do.

Here is the dialogue that followed (see Schoenfeld, 1998, 1999, for a full transcript and analyses):

1. MN: OK ($\frac{x^5}{x^5}$). What did you get here?
2. Ss: [overlapping] x to the zero. Nothing. Just zero. One.
3. MN: [pointing to various students] One. x to the zero.
4. OK. [Writes $\frac{x^5}{x^5} = \underline{\hspace{2em}}$ on the board.] Let's work this out. Let's work this one out. [pointing to numerator] This says how many x's do we have?
5. Ss: Five.
6. [Writes $\frac{x^5}{x^5} = \frac{xxxxxx}{}$]
7. MN: And how many x's are on the bottom?
8. Ss: Five.
9. MN: [Writes $\frac{x^5}{x^5} = \frac{xxxxx}{xxxxx}$] OK.
10. MN: So what do I do?
11. Ss: Cancel.
12. MN: [Cancels with a series of slashes.] $\frac{\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}}{\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}}$
13. MN: So what am I left with?
14. Ss: [overlapping] x. Zero. x. [inaudible.]
15. MN: There's nothing here? Is there zero here?
16. Ss: Yes. No. Zero over zero. Nada
17. MN: [holding up hand in gesture to wait] Just a second. Just a second.
18. MN: [moves to blank section of board and writes] What if we had 5 over 5? What's that equal?
19. Ss: One.
20. MN: But we, I just canceled them.
21. S4: Nothing.
22. S5: But there's a 1 there.
23. MN: [Points to the slashed 5/5] But there's a 1 there. [Points to the expression:] $\frac{\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}}{\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}\,\cancel{x}}$
Is there a 1 there?

24. Ss: [overlapping] Yes. No.
25. MN: Yeah there is.
26. S6: There ain't no 1 there.
27. MN: [visibly slumps at the board]

Why did MN react the way he did? We know that he knows the relevant mathematics and is more than capable of explaining that when $x \neq 0$, $\frac{xxxxx}{xxxxx} = \left(\frac{x}{x}\right)^5 = (1)^5 = 1$. Instead he slumped at the board, seemingly without resources.

Case 2

JM, an experienced teacher, is in the middle of an introductory lesson in his high school physics class. The challenge is for the class to settle on the "best value" for the width of a table, given that the values that eight different students had obtained for the width of the table were 106.8; 107.0; 107.0; 107.5; 107.0; 107.0; 106.5; and 106.0.

The class, aware that "outliers" are sometimes dismissed as data, has considered whether or not to include all of the numbers in their computation of the best value. They have gone on to consider how best to combine the numbers. In a conversation with a fair amount of give and take, they considered and defined the average, median, and mode. At that point, JM once again asks "Anybody think of another way of giving a best value?" A student says,

> This is a little complicated but I mean it might work. If you see that 107 shows up 4 times, you give it a coefficient of 4, and then 107.5 only shows up one time, you give it a coefficient of 1, you add all those up and then you divide by the number of coefficients you have.

A teacher who encounters this kind of situation might respond, "Interesting question. Let's talk about that after class" (or " . . . tomorrow in class"). Or the teacher might praise the student for her insight and briefly explain the idea of a weighted average. Or the teacher might open this comment up for discussion. Each approach has costs and some benefits in terms of time, lesson flow, content coverage, student enfranchisement, and so on. Can one say what JM is likely to do and why?

Case 3

DB is leading her third-grade class in a discussion of the previous day's meeting with fourth-graders, in which there had been an extended conversation about the nature of zero—is it even, odd, or "special?" Her purpose is not to focus on the mathematics at this point but to encourage the students to reflect on their learning—how do they think about their own understandings and how were those understandings affected by what they heard at the meeting? In a sequence of interactions with students, DB notes that one student was prompted to think more deeply about the content by the previous day's discussion; she asks another what she's going to do given that she was made less certain of her understanding by some things that were said. After two students have a rather uncomfortable exchange about the rationale for a claim that zero is even, she emphasizes that these are difficult matters that take some sorting through. Her emphasis has consistently been to go "meta"—to reflect on the process of understanding. However, Benny then makes the following comment about the nature of zero:

> "Um, first I said that um, zero was even but then I guess I revised so that zero, I think, is special because um, I—um, even numbers, like they make even numbers; like 2, um, 2 makes 4, and 4 is an even number; and 4 makes 8; 8 is an even number; and um, like that. And, and go on like that and like 1 plus 1 and go on adding the same numbers with the same numbers. And so I, I think zero's special,"

DB now violates her own "meta" rule and asks about content:

> "Can I ask you a question about what you just said? And then I'll ask people for more comments about the meeting. Were you saying that when you put even numbers together, you get another even number, or were you saying that all even numbers are made up of even numbers?"

The result is a significant detour from DB's announced agenda of reflecting on the previous day's meeting. Might there be a plausible explanation for this seemingly odd behavior?

Here is my explanation of what happened in the three cases. My main purpose in offering these explanations is to explain how the key terms in our models of teaching help to provide explanations. I then go on to characterize the modeling enterprise more generally.

Case 1: MN

MN knows the relevant mathematics and is perfectly capable of explaining it. Yet he slumps defeatedly at the blackboard when his students justifiably demonstrate confusion upon being asked to reduce the fraction $\frac{x^5}{x^5}$ Why?

The answer is a combination of MN's *beliefs* and *knowledge.* MN had a number of beliefs about teaching, including the belief that "telling" by itself is not a good thing. He believed that his role as a teacher was to work with ideas and explanations generated by his students and to help the students refine them. That is, he felt that he had license to elaborate and clarify only when the core idea has been expressed by a student. This is what he had done with the first two problems, $\frac{m^6}{m^2}$ and $\frac{x^3y^7}{x^2y^6}$. When a student said "subtract" in response to MN's question about how he got the answer, the reference to subtraction gave MN "permission" to elaborate and to offer a fuller explanation of how to solve the problem. However, the discussion of $\frac{x^5}{x^5}$ never got to that point—the students never gave him the raw material upon which he felt he could elaborate. Lacking the pedagogical knowledge to induce the students to generate what he wanted and believing that it was inappropriate for him to introduce that material himself, he was completely stymied.

Case 2: JM

The key to knowing how JM will respond to the student's comment is to understand his *goals and beliefs* as well as the tools he has for implementing them (his pedagogical *knowledge*). As explained in Schoenfeld (1998),

> "JM conceives of physics as a sense-making activity, and he believes that students' classroom experience with physics should support the development of students' abilities as sense-makers. He encourages open and free discussion in class and has the class spend a great deal of time thinking things through. In order for students to do the kind of sense-making he would like, they need to engage with the material and not simply be told what works. Hence JM works to create a discourse community that involves small-group work among students, and he structures teacher–student exchanges in a way that minimizes "telling." As part of his commitment to student inquiry, JM values and places very high priority on pursuing substantive ideas raised by students—indeed, pursuing important ideas raised by students will tend to get higher priority from JM than following a predetermined lesson plan." (p. 48)

One can model JM's *decision-making* by performing cost–benefit analyses, similar to those used in economics to evaluate "subjective expected value," for each of the options outlined in case 2. For example, shunting the student's question aside would keep the lesson on track, which is a plus, but it would mean losing an important early opportunity to demonstrate JM's commitments to sense making and student enfranchisement—a big minus. In contrast, pursuing the student's question at length demonstrates JM's values and helps to establish the classroom norms JM wishes to establish. These benefits (from JM's point of view—again, these are personal and subjective judgments) far outweigh the costs of lost lesson time. Thus it is clear that JM will follow up on the student's comments. That is the choice our model of JM makes.

In addition, when one knows enough about JM's teaching knowledge, one can predict not only that he will pursue the student's comment but also how he will do so. An important part of JM's pedagogical repertoire consists of what he calls "reflective tosses," questions designed to elicit clarifications from the students about the issues under discussion. Having made the decision to follow up on the student's comment, it is likely that JM will ask a series of questions intended to make sure the students understand his proposal and that they work through whether or not it is correct.

Case 3: DB

My focus here is on understanding an unusual instance of a teacher's *decision making.* Numerous people watching the tape of DB's class have commented on what they perceive to be the strangeness of her decision to ask Benny if (in our language) he means that every even number is double another even number, or that doubling any even number results in another even number. This question seems mathematically tangential, and pursuing it appears to undermine DB's announced "reflection" agenda. Could there be a reasonable explanation for this? If no, then my research agenda of modeling teachers' actions and decisions could be seriously compromised—it implies that some percentage of experienced teachers' decisions is simply "random."

The key to understanding DB's decision making in this case is having a deeper understanding of her *knowledge, goals,* and *beliefs.* Like JM, DB is committed to developing communities of inquiry in her mathematics classrooms. This means enfranchising the students, and developing appropriate norms of interaction. It means monitoring and assessing what the students know, because the directions of their conversations (and their utility!) depend very much on the students' understandings. DB's notes for the class indicate that the larger frame for this particular class was a unit on "number theory." The class had been exploring the properties of even and odd numbers. Some students had observed that whenever they added two even numbers or two odd numbers, the result was always even. DB's plan for the bulk of the day's lesson was to have the students explore those conjectures, with her facilitation. What did Benny understand? If he thought that doubling any number resulted in an even number, that was one thing. If he (and at least one other student, Bernadette, who had said something similar) thought that every even number was double another even number, she would face a very different pedagogical task. A close look at what DB says ("Can I ask you a question about what you just said? And then I'll ask people for more comments about the meeting.") indicates that she does not intend to forsake her reflection agenda. Rather, she is putting that agenda on hold for a short while in order to gather information that is important for the success of the main part of the day's lesson. Thus, what appears on the surface to be a somewhat capricious decision, which runs the risk of derailing a reflective conversation she has taken pains to set in motion, can be understood—once one has adequate information about her knowledge and intentions—as a plausibly motivated attempt to gather information necessary for the success of the major part of her planned lesson.

A THEORY OF TEACHING-IN-CONTEXT, PART 2: A BRIEF DESCRIPTION OF THE MODELING ENTERPRISE

The preceding section illustrates some key aspects of the models of teaching, with a focus on explaining the three teachers' decision making. Here, I wish to convey in more fine-grained detail the flavor of the full analyses. Thus I recap the first part of the analysis of the lesson that gave rise to Case 3 above. In what follows I provide a brief analysis of the first nine turns of dialogue of that day's lesson. The full analysis is given in Schoenfeld (2008).

As noted above, DB had been teaching a unit in which her students explored the properties of even and odd whole numbers. She had taught a similar unit the previous year. On the day before this lesson, DB's current class had met with last year's class to discuss even numbers, odd numbers, and zero. An issue had arisen as to whether zero was even, odd, or "special"; but many other issues were also on DB's mind. Overall, DB's agenda for her students is expansive. Her goals for students include not only mathematical content but also their development as learners, as members of the classroom community, and more (see Ball, Lewis and Thames, 2008, for a discussion of teachers' work; see Lampert, 2001, for a discussion of the multiple goals teachers have for their students). Although the main focus of DB's lesson would be a continuation of the discussion of even and odd numbers, DB wanted to start the lesson by reflecting on the previous day's experiences.

The lesson occurs midyear, which means that norms for discussion in this classroom are well established. Moreover, as an experienced teacher, DB herself has a large repertoire of classroom routines. One such routine, a "flexible, interruptable routine for discussing a topic" (see Figure 9.1), is employed at the beginning of this lesson. In brief, the routine works as follows. DB raises an issue, asking students for comment. When a student responds, DB must decide whether the student's comment is "on topic" or not and whether to pursue it. (If she decides not to pursue it, she will steer the conversation back to the original issue.) When the student's comment is on topic, DB decides whether it calls for clarification or expansion and acts accordingly. After the student's comment has been worked through (clarified, elaborated, etc.), DB decides whether she wishes to expand on it or summarize the gist of her exchange with the student for the class. Then she will either invite further comments or, if she feels the discussion has run its natural course, turn to another issue.

Of course, DB's decision making during the implementation of this or any other routine depends on her short- and long-term goals, her beliefs about what is important, and other contingencies (for example, how much time she thinks she can profitably devote to the routine).

Figure 9.2 shows a parsing of the first nine turns in the lesson. At the descriptive/analytic level, Figure 9.2 can be read as follows. The first three columns of Figure 9.2 are a transcript containing, respectively, the turn number, the name of the person speaking, and what the speaker says. These are followed by two levels of parsing. The lesson is broken into lesson segments called "episodes," which cohere on phenomenological grounds. These episodes are further decomposed into subepisodes and so on if needed. (In analyses of larger segments of a lesson, there may be a half dozen levels of parsing.) In DB's lesson as a whole, the first major episode consists of the full classroom discussion of the previous day's meeting. The discussion of the previous day's meeting consists of some brief preliminary comments to get the class under way, followed by five full iterations of the routine described in Figure 9.1. (Only the first iteration is shown in Figure 9.2.)

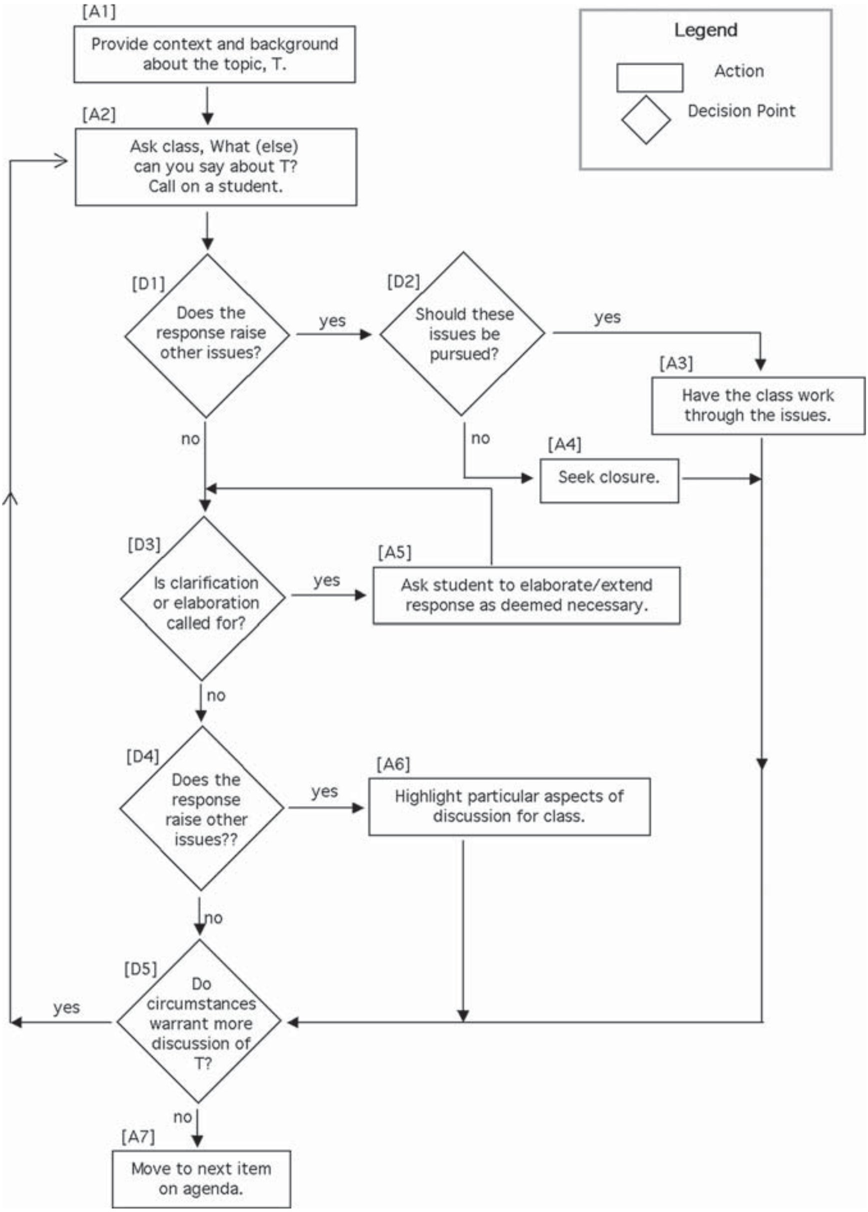


Figure 9.1 A complex routine for soliciting and working with student ideas.

Here, DB's decision making is straightforward. She entered the classroom with the goal of debriefing the students on the previous day's meeting. The routine characterized in Figure 9.1 is a readily accessible part of her knowledge base, and she moved to implement it in turn 1 of the transcript. Once the routine is begun, it calls for a series of contingent decisions, which she made in turns 3, 5, 7, and 9. DB's decisions and the knowledge base that supports them are described in the two rightmost columns of Figure 9.2.

The columns in the center of Figure 9.2 show the activation levels of a range of DB's relevant beliefs and goals as the lesson unfolded. For reasons of space the full list of goals and beliefs is not included here, and I omit a discussion of some of DB's goals and beliefs, for example, DB's very broad (overarching) goals A, B, and C and her general learning beliefs L1 and L2. Some of the most highly activated goals during this lesson segment are outlined here.

Major content and learning goals for this lesson segment

D. Debrief students on the previous day's meeting (and let the discussion take its course as long as this course is consistent with other high-priority goals).

E. Focus on how students understand their own mathematical understanding and learning in the context of (and using as data) the previous day's meeting.

Local goals (in the context of implementing the routine in Figure 9.1)

F. Have a student clarify, elaborate, or extend what he or she has said (turns 3–5).

G. Highlight a student's statement (perhaps transforming it or adapting it) in service of top-level instructional goals (turn 7).

Similarly, some of the most highly activated beliefs during this lesson segment are:

Beliefs about mathematics learning

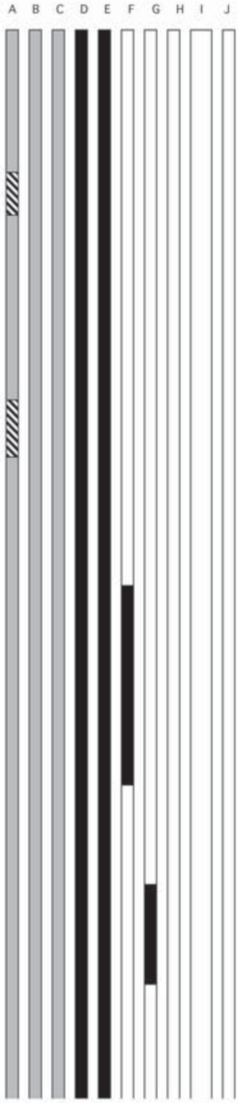
(L3) Mathematics learning includes learning to learn. This includes reflecting on what one knows and reflecting on how one learns.

(L4) Mathematics learning occurs through participating in mathematical discourse. This includes generating mathematical

Line	Who	Transcript	First Level of Parsing	Second Level of Parsing
1	Ball	[Preliminaries] [A] Okay. A few delays, but I think we're ready to start now.	[Preliminaries] Getting the class started	
		[1] [B] I'd like to open, open the discussion today with um – I have a few questions about the meeting yesterday that I'd like to ask.	[1.1] Initiation of debriefing routine, and (as planned) execution of routine through one full cycle.	[1.1.1] Establishing Context and background (Action A1)
		[C] So, to begin with, I would just like everybody to put pens down, there's nothing to take notes about or do right now.		Classroom management
		[D] But I'd like you to be thinking back to yesterday and to the meeting that we had on even and odd numbers and zero.	See sections 4.2 and 4.3 forprose and flow chart descriptions of this routine.	
		[E] And I have a few questions. First – my first question is, I'd just like to hear some comments about what you thought about the meeting, what you noticed about the meeting, what you learned at the meeting, just what kinds of comments you have about yesterday's meeting?	(Decisions listed in this column are represented in Figure 3.)	[1.1.2] Posing the main question (Action A2)
		[F] And could you listen to one another's comments, so that we can um, benefit from what other people say? See what y– what you think about other people's comments?		Classroom management
		[G] Sheena, do you want to start?		Calling on student
2	Sheena	I– I– I liked it because, well, I like talking to other classes and, and when you talk to other classes sometimes it helps.	[D1] = no [D3] = yes	[1.1.3] Student Response
3	Ball	In what way?		[1.1.4] Request Clarification (Action A5)
4	Sheena	It helps you to understand a little bit more.	[D3] = yes	[1.1.5] Student Response
5	Ball	Was there an example of something yesterday that you understood a little bit more during the meeting?		[1.1.6] Request Clarification (Action A5)
6	Sheena	Well, I didn't think that zero was – zero, um – even or odd until yesterday they said that it could be even because of the ones on each side is odd, so that couldn't be odd. So that helped me understand it.	[D4] = yes	[1.1.7] Student Response
7	Ball	Hmm. So y– So you thought about something that came up in the meeting that you hadn't thought about before? Okay.		[1.1.8] Abstract and reframe student comment (Action A6)
8	Sheena	(Nods in agreement.)	[D5] = yes	(Cycle ends)
9	Ball	Other people's comments? Sean?	[1.2] 2nd cycle of routine	[1.2.1] initiation of 2nd cycle (Action A2)

Goals:

See narrative for descriptions of relevant goals
Goals A, B, and C are overarching goals
Goals D and E are major content goals
Goals F,G,H,I, and J are local goals



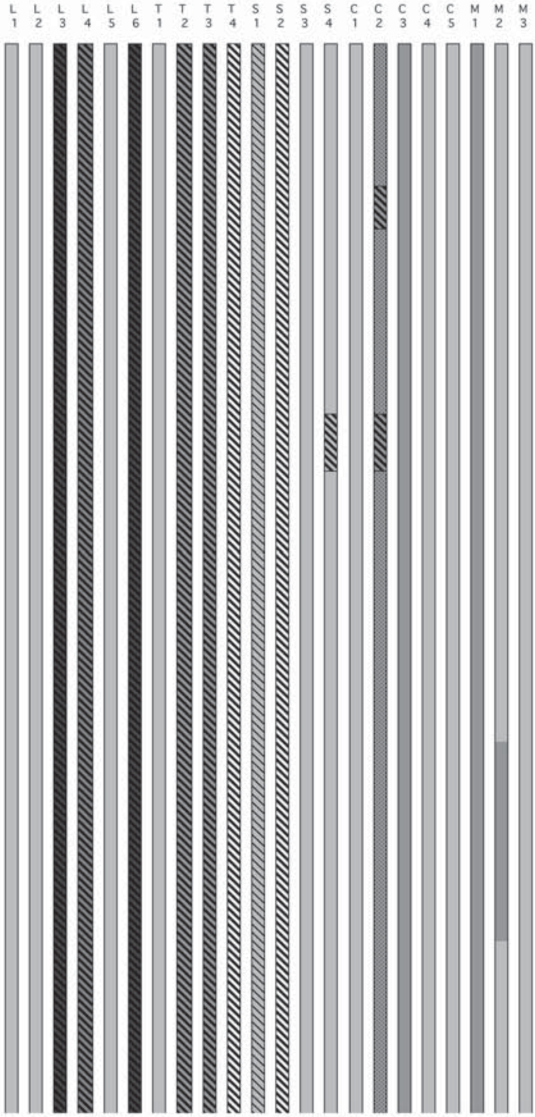
Legend

Latent, ready to be activated
Present on an ongoing basis
Present throughout at a strong level

Figure 9.2 A parsing of the first nine lines of student-teacher dialogue in DB's lesson.

Beliefs:

See narrative for descriptions of relevant beliefs
Beliefs L1-L6 are about learning
Beliefs T1-T4 are about teaching
Beliefs S1-S4 are about students
Beliefs C1-C5 are about classroom environments
Beliefs M1-M3 are about mathematics



Relevant Knowledge

Previous days' converations
Experience of previous year's conversations
Individual students' backgrounds
Her thoughts about student understandings of the topic and about the relevant mathematics (evenness, oddness, the possibly "special" nature of 0).

Ball has ready access to the debriefing routine - it's a standard part of her repertoire.

Ball has a general scaffolding routine that she uses to get students to clarify/ elaborate their thoughts. This, combined with her specific knowledge of Sheena, suggests how much scaffolding is necessary.

Part of Ball's knowledge is the ability to abstract the content of conversations with students in the service of her goals, and to present that content at an appropriate level to her students

Decision-Making

Ball begins the class with the intention of "debriefing" the students about the previous day's meeting.

She plans to implement the "flexible, interruptable" debriefing routine described in Figure 1.

Sheena's comment is on target ([D1] = no), so Ball will pursue it. It is vague ([D3]=yes), so Ball will provide scaffolding to help her clarify what she means to say.

[D3] = yes.
Sheena's statement is still vague, so Ball will make a 2nd move for clarification, using a standard "give me an example" prompt in response to a vague statement

This time Sheena's answer is clear ([D3] = no) so Ball moves on.

[D4] = yes. Ball reframes what Sheena has said at a meta-level: one hears new things at meetings.

[D5] = yes
Ball pursues the debriefing routine.

Present but temporarily on hold | At medium activation level | At high activation level | At the highest activation level ideas, listening to others' mathematical ideas, and critically examining mathematical ideas, as individuals and collectively;

(L6) Mathematics learning is facilitated by articulating one's thoughts, orally and in writing.

Beliefs about teaching

(T1) Teaching includes establishing, building, and maintaining a classroom community that enfranchises students and supports sense making.

(T2) Teaching includes probing and making sense of students' understandings and beliefs.

On the one hand, a characterization of the type given in Figure 9.2 can be read as a detailed description and analysis of a teacher's actions. On the other hand, Figure 9.2 can also be used as the basis for a detailed cognitive model of DB's teaching (of this lesson). Imagine, for example, that one has just gotten to turn 4 in the classroom. One can ask, what is DB likely to do next? The answer in this case is as follows.

She is in the middle of the routine described in Figure 9.1. Sheena's comment in turn 4 brings DB to decision point [D3] in Figure 9.1. Here [D3] = yes, because Sheena's statement is still vague. Thus DB will make a second move for clarification, employing a standard "give me an example" prompt in response to Sheena's vague statement.

In this way, the detailed characterization of DB's knowledge, goals, beliefs, and decision making allows one to construct a model of DB's teaching. In some straightforward instances, such as this one, the model allows one to predict with a high degree of confidence what DB is likely to do. In other instances, such as the one described in case 3, no such prediction is possible—but the model can be used to generate a set of plausible alternatives from which DB is likely to choose.

More generally, the kind of model described here is an instance of the theory of teaching in context described in Schoenfeld (1998, 2008, in press). The core idea of the theory is that in-the-moment decision making can be characterized as a function of an individual's knowledge, goals, and beliefs. This is the case not only for teaching but for many of the tasks in which we engage in general. The description in Exhibit 9.1 is at its most general, but it is easy to see how it applies in the classroom.

Exhibit 9.1

HOW THINGS WORK, IN OUTLINE.

- An individual enters into a particular context with a specific body of knowledge, goals, and orientations (beliefs, dispositions, values, preferences, etc.).
- The individual orients to the situation. Certain pieces of information and knowledge become salient and are activated.
- Goals are established (or reinforced if they preexisted).
- Decisions are made, consciously or unconsciously, in pursuit of the higher-order goals.
  - If the situation is familiar, then the process may be relatively automatic, where the action(s) taken are in essence the access and implementation of scripts, frames, routines, or schemata.
  - If the situation is not familiar or there is something nonroutine about it, then decision making proceeds via an internal calculus that can be modeled by (i.e., is consistent with the results of) the subjective expected values of available options, given the orientations of the individual.
- Implementation begins.
- Monitoring (whether effective or not) takes place on an ongoing basis.
- This process is iterative, down to the level of individual utterances or actions:
  - Routines aimed at particular goals have subroutines, which have their own subgoals.
  - If a subgoal is satisfied, the individual proceeds to another goal or subgoal.
  - If a goal is achieved, new goals kick in via decision making.
  - If the process is interrupted or things do not seem to be going well, decision making kicks into action once again. This may or may not result in a change of goals and/or the pathways used to try to achieve them.

ISSUES OF THEORY, MODELS, AND APPLICATIONS

I now turn to the question of why one might engage in the kinds of intense modeling activities described in this chapter. There are, I believe, both theoretical and practical reasons for this. On the theoretical side, two major questions are (a) how do you know that your theoretical ideas are accurate and (b) how do you know they are general?

A significant issue for those proposing explanations in the social and behavioral sciences is the danger of "ad hocism"—a tendency to establish temporary, chiefly improvisational explanations of actions taken by individuals. Taken one at a time, these explanations may seem plausible; but in toto the whole of the explanation of an individual's behavior may be less than the sum of its parts. A theory commits one to certain classes of explanation—in the case of the theory of teaching (or acting) in context, to explanations that employ only the constructs of knowledge, goals, beliefs, and a particular form of decision making. But a major test of an abstract theory is its ability to accurately model a wide range of instances ostensibly governed by the theory. For example, Copernican theory required so many "fixes" (epicycles upon epicycles) that there was good reason to question its adequacy. In contrast, Newtonian gravitational theory provides reasonably accurate models of our solar system. One inputs the masses, velocities, accelerations of the planets and the sun and applies the inverse-square gravitational law. The resulting models corresponds fairly well to the actual positions of the planets, giving one some confidence in the theory. Similarly, by building models of teaching that depend on (and are limited to) analyses of the teachers' knowledge, goals, and beliefs we have a way of guarding against ad hocism; if the modeling of new exchanges consistently calls for modifications, then the explanatory power of the theory is in question.

Similarly, there are issues related to the scope and generality of a theory.[3] Just how broadly does a theory apply? One test is to see how many different classes of phenomena can be modeled. Newtonian theory—a grand theory of physics if there ever was one—would not be that grand if it applied only to our solar system. But in fact it applies reasonably well to a wide range of physical phenomena.[4] Thus, my strategy for testing the theory involved selecting a range of cases. These included a beginning high school teacher (MN) teaching rather traditional lesson content; an experienced high school teacher (JM) teaching an innovative lesson of his own design; a university mathematician (myself) teaching a highly interactive problem solving course; and an experienced third-grade teacher (DB) teaching a lesson whose agenda evolved substantially as a result of comments made by the students. As the space of modeled examples grew, there was increasing reason to believe in the robustness and generality of the theory.

At this point there is reason to believe that the model applies to just about any instance of classroom teaching (as long as the teacher is not random; one cannot model random behavior, of course). The largest challenge came in 1998. I reported (in Schoenfeld, 1998) that although I had managed to model MN's and JM's lessons, I had thus far been unable to construct a model of DB's lesson. That lesson was a challenge for two reasons. First, unlike the teacher-directed lessons I had succeeded in modeling up to that point, the agenda for DB's lesson agenda was coconstructed with her students as the lesson evolved. Second, third-graders differ substantially from high school students; perhaps differences between the pedagogies in high school and third grade would be so large that the latter could not be modeled using the techniques that had enabled me to model the former. Had I not succeeded in modeling DB's lesson, there would have been clear evidence of the limitations of the theoretical approach. But the success in modeling DB (Schoenfeld, 2008) suggests how broadly the theory does apply. Moreover, there is reason to believe (Schoenfeld, in press) that people's behavior in a wide range of domains (shopping, cooking, automobile mechanics, medical practice) can be similarly modeled.

There are a number of other issues one could raise. For example, are there any competing theories, and how would one compare two theories that claimed to explain the same thing? To my knowledge, there are no competitors at this level of detail—that is, I know of no other theoretical/empirical approach that sets out to explain and predict, on a line-by-line basis, what teachers do. Should such a theory be proposed, a comparative test of the two theories would be simple: of the two, which comes closer to accounting for the empirical evidence? In the research papers that describe the theory of teaching in context, I provide full transcripts of the classroom sessions that we model. One reason I do this is to provide readers with a large data base against which to check our interpretations of teacher actions. But another is that I invite alternative models. The only constraint on alternative models is that they cannot be ad hoc, providing one reason for this behavior, one reason for that behavior, etc. There is a form of internal consistency built into our models in that we have to specify the architecture, load in the parameters, and then step back and run the model. Only models of that type are legitimate competitors.

More at a metalevel, one can question the theoretical entities that are core to the theory. Does one believe in goals, for example? Does one believe in beliefs? That people have them, are guided by them? Here one must be very careful. Models are just that—models. One never knows for certain whether people have particular beliefs, or that their beliefs (or goals, or knowledge, or decision making) really function in particular ways. All one can say is that a model with those beliefs, goals, and knowledge behaves in the same way the person does. It stands to reason that if something walks like a duck and talks like a duck . . . it is probably a pretty good model of a duck.

Now on to the issue of applications: is this kind of work potentially useful, or is it "just" theoretical? From my perspective there is a range of potential applications, some of which we have begun to explore. By way of analogy, my earlier work in problem solving (e.g., Schoenfeld, 1985) provided descriptions of what "counts" when people attempt to solve mathematics problems. That theoretical account had an instructional counterpart: knowing that heuristic strategies needed to be described at a particular level of detail, that monitoring and self-regulation are important, and that students' beliefs shaped the approaches they took to problems, I was able to construct courses that focused on these aspects of problem solving and that enabled students to become more effective problem solvers.

I believe that similar attempts have the potential to enhance the preparation and professional development of teachers. For example, the classroom routine characterized in Figure 9.1 is more general than one might imagine. In style, JM, DB, and I are three very different teachers. I had characterized (in Schoenfeld, 1998) a teaching routine used by JM in which he elicited comments and suggestions from students. Long after I began the analysis of DB's lesson, I came to realize that she used the routine captured in Figure 9.1—and only then did I realize that this was, at that level of abstraction, the same routine that I had ascribed to JM! It turns out that I use the same routine in my problem-solving courses. This routine, having been abstracted, can now be taught (see Schoenfeld, 2002).

Second, we know that beliefs and various aspects of metacognition play a fundamental (and often unconscious) role in shaping teachers' decision-making. In a preliminary study, Arcavi and I (Arcavi & Schoenfeld, 2008) structured a set of experiences to introduce practicing teachers to the idea that one's beliefs shape one's practices and to induce the teachers to reflect on their own beliefs. We began by having teachers watch videotapes of teachers from other countries (e.g., the TIMSS videotapes of Japan; see Stigler & Hiebert, 1999). The teachers' reactions upon seeing practices that were nonstandard (for them) was "We couldn't do that here. They can do it because their students are different." Shifting the question to "Why might a teacher engage in that practice? What might a teacher think or believe about teaching or learning, that would make that practice seem a reasonable thing to do?" opened up the issue of beliefs in a nonthreatening way, and allowed the teachers to think about how their own beliefs might shape their practices. The results, though preliminary, are promising.

A third potential application concerns the possible diagnostic use of theoretical frameworks such as the one discussed here. At present the modeling of any particular lesson—at the level of analysis used for research—is costly in terms of time and effort. But that is partly because our research goal was to model individuals in detail. To return to the problem-solving analogy, there now exist tests of problem solving that provide evidence of what important concepts and processes students have and have not mastered. These tests provide useful information both at individual student level and for possible curriculum change at the district level (see, e.g., Foster, Noyce, & Spiegel, 2007). These inquiries into student knowledge are not conducted at the level of detail that the original research on mathematical thinking and learning was conducted, of course—but they provide useful information that can be gathered in relatively short order. One can imagine similar observational assessments of teachers. The goal of the research has been to identify the key dimensions of teaching performance. But once those key dimensions have been established, it may be possible to get relatively quick "readings" of teachers' strengths and weaknesses and then to target professional development accordingly. Indeed, this kind of assessment might be used to see if there are typical developmental trajectories for teachers, and to help identify particularly useful interventions at particular points in teachers' careers. As the knowledge base on teaching gets solidified, the potential for applications increases.

LIMITATIONS AND NEXT STEPS

I should stress that the current work provides a theory of teaching (or in its most general form, a theory of acting) *in the moment.* It says, in essence, that if we know a person's knowledge, goals, and beliefs at a particular time, then we can model that person's decision-making at that particular time. A major category of intellectual behavior that the current theory does not address is the growth and change of people's understanding—that is, learning. As a result of teaching a particular lesson, a teacher may develop refined understandings of his or her students, the mathematical content, and ways to implement the lesson that has just been taught. The current theoretical model does not contain a mechanism for incorporating knowledge growth and change. As such, it provides a mechanism for creating only "snapshots" of performance. A significant (and I believe possible) next theoretical step is to attend to issues of learning and model professional growth as well as performance.

A second issue that needs to be addressed is that of perspective. The current theory is a theory of individual cognition. It views classroom events from the perspective of one important participant, that of the teacher. Thus, all the individuals and events in the class are seen through the (quite possibly biased) lens of the teacher's understanding. Of course, this is how the teacher sees things—but it is only one slice of classroom reality. Every participant has his or her own version of that reality, and they are all different. More importantly, "classroom reality" is much more than the sum of the views of the individuals. The question of how all of those individuals coconstruct the classroom community and live in dialectic interaction with it is one of the great theoretical challenges of our time. It may be some time before we make significant headway on that problem. If, however, one recalls that the field is less than a half-century old and one contrasts our current state of knowledge with what was known in the 1960s and 1970s, there is tremendous reason for optimism.

NOTES

1. Of course, a particular tutor can privilege content over affect, or vice versa. The idea behind an "architecture" is that it specifies how a family of models works. Then, when you build a model of a particular tutor, you specify the characteristics of that tutor as parameters in the more general architecture. Tutor A might have a high threshold for student boredom or unhappiness and persevere with the mathematics, while tutor B might have a comparably low threshold and seek to assure or find alternative activities much earlier. The models of tutors A and B would have the same structure, but their "reaction thresholds" would be set differently.
2. There are, presumably, social niceties that get taken care of as precursors to the tutoring interactions. These may feed into affective decision making during the tutoring session.
3. For an extended discussion of these and other dimensions of theoretical and methodological importance, see Schoenfeld, 2007.
4. I say "reasonably well" because Newton's theory was supplanted at the cosmic scale of interactions by Einstein's theory of relativity. But any broad theory that survives for hundreds of years certainly deserves the label *grand*.

REFERENCES

Arcavi, A. A., & Schoenfeld, A. H. (1992). Mathematics tutoring through a constructivist lens: The challenges of sense-making. *Journal of Mathematical behavior, 11*(4), 321–336.

Arcavi, A. A., & Schoenfeld, A. H. (2008). Using the unfamiliar to problematize the familiar. *Canadian Journal of Science, Mathematics and Technology Education, 8*(3), 280–295.

Ball, D., Lewis, J., & Thames, M. (2008). *Making mathematics work in school.* In A. H. Schoenfeld, (Ed.) *A study of teaching: Multiple lenses, multiple views.* Journal for Research in Mathematics Education monograph series (No. 14, pp. 13–44). Reston, VA: National Council of Teachers of Mathematics.

Foster, D., Noyce, P., & Spiegel, S. (2007). When assessment guides instruction: Silicon valley's Mathematics Assessment Collaborative. In A. H. Schoenfeld (Ed.), *Assessing Mathematical proficiency* (pp. 137–154). Cambridge, UK: Cambridge University Press.

Lampert, M. (2001) *Teaching problems and the problems of teaching.* New Haven, CT: Yale University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Newell, A., & Simon, H. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Pólya, G. (1945). *How to solve it.* Princeton, NJ: Princeton University Press.

Pólya, G. (1954). *Mathematics and plausible reasoning* (Vol. 1, *Induction and analogy in mathematics;* Vol. 2, *Patterns of plausible inference*). Princeton, NJ: Princeton University Press.

Pólya, G. (1962). *Mathematical discovery* (Vol. 1). Princeton, NJ: Princeton University Press. (Combined paperback edition, 1981. New York: Wiley)

Pólya, G. (1965). *Mathematical discovery* (Vol. 2). Princeton, NJ: Princeton University Press. (Combined paperback edition, 1981. New York: Wiley)

Schoenfeld, A. H. (1985). *Mathematical problem solving.* Orlando, FL: Academic Press.

Schoenfeld, A. H. (1998). Toward a theory of teaching-in-context. *Issues in Education, 4*(1), 1–94.

Schoenfeld, A. H. (Special Issue Ed.) (1999). Examining the complexity of teaching. *Journal of Mathematical Behavior (special issue), 18*(3).

Schoenfeld, A. H. (2000) Purposes and methods of research in mathematics education. *Notices of the American Mathematical Society, 47*(6), 2–10.

Schoenfeld, A. H. (2002). A highly interactive discourse structure. In J. Brophy (Ed.), *Advances in research on teaching. Vol. 9, Social constructivist teaching: Its affordances and constraints* (pp. 131–170). New York: Elsevier.

Schoenfeld, A. H. (2007). Method. In F. Lester (Ed.), *Handbook of research on mathematics teaching and learning* (2nd ed., pp. 69–107). Charlotte, NC: Information Age Publishing.

Schoenfeld, A. H. (2008). On modeling teachers' in-the-moment decision-making. In A. H. Schoenfeld (Ed.), *A study of teaching: Multiple lenses, multiple views*. Journal for Research in Mathematics Education monograph series (No. 14, pp. 45–96). Reston, VA: National Council of Teachers of Mathematics.

Schoenfeld, A. H. (in press). *How we think.* New York: Routledge.

Schoenfeld, A. H., Gamoran, M., Kessel, C., Leonard, M., Orbach, R., & Arcavi, A. A. (1992). Toward a comprehensive model of human tutoring in complex subject matter domains. *Journal of Mathematical Behavior, 11*(4), 293–320.

Stigler, J., & Hiebert, J. (1999). *The teaching gap.* New York: Free Press.

This page intentionally left blank

10 Uncovering the Role of Culture in Learning, Development, and Education

CARRIE ROTHSTEIN-FISCH, PATRICIA M. GREENFIELD, ELISE TRUMBULL, HEIDI KELLER, AND BLANCA QUIROZ

> Each child in kindergarten is making a book called *All About Me.* The teacher asks each child to say what is special about himself or herself, writing down these comments, with space for the child to draw a picture corresponding to each comment. One boy responds with "my brother is good at soccer and my father is good at cooking." The teacher keeps asking the boy to think about qualities of himself, prompting him: "This is all about you. Are you smart? Yes, of course you are smart, so let's say you are special because you are smart." In the end, the boy's book contained drawings with these sentences "I am special because I am smart. I am special because I am strong. I am special because I am handsome." (Zepeda, Gonzalez-Mena, Rothstein-Fisch, & Trumbull, 2006, p. 19)

The practice of making a book called *All About Me* is a familiar activity in many early childhood classrooms. But there may be more to this activity than what meets the eye. The teacher is likely to be thinking that identifying concepts of self will promote the child's sense of self-esteem and individuality. She may also value the child's burgeoning literacy skills. Yet, the child may be experiencing something very different. He may believe that the teacher does not like his family, and thus she does not like him. Bragging about himself may make him feel very uncomfortable. Ultimately, he may be inclined to think that his values and ideas do not matter, thwarting his concept of self—exactly the opposite of the teacher's goal (Zepeda et al., 2006).

The example of *All About Me* calls attention to classroom practices that, though well-intended, may be at odds with learning, eventually leading to negative feelings about school altogether. The good intentions of the teacher and the compliant but uncomfortable boy are likely to be operating with two conflicting sets of values, each invisible to the other. The teacher's goals are representative of the cultural value of *individualism,* the characteristic value of mainstream United States. In contrast, the boy's discomfort at being isolated from his family is characteristic of the cultural value of *collectivism,* the value system of many immigrant children and families. Individualism and collectivism have emerged as powerful cultural models that tie together cultural conceptions of learning and development, drawing on theory and research in developmental psychology (Greenfield & Bruner, 1966) and anthropology (Whiting & Whiting, 1975). These two idealized developmental pathways emphasize different goals for development and learning. Individualism emphasizes individual identity, independence, self-fulfillment, and standing out. Collectivism emphasizes group identity, interdependence, social responsibility, and fitting in (Greenfield, Keller, Fuligni, & Maynard, 2003).

Each pathway is situated in a broader sociocultural system (Keller, 1997, 2003). The individualistic pathway arises as an adaptation to a complex, urban, wealthy environment featuring a well-developed system of formal education and advanced technology. The collectivistic pathway arises as an adaptation to a small-scale, face-to-face village environment based on a subsistence economy and informal education (Greenfield, Trumbull, et al., 2006). Economic conditions and political persecution tend to incorporate people from the second kind of society into the first (Greenfield, 2009). When this happens, children and their families are exposed to two contrasting and often conflicting socializing forces that are very relevant to the care and education of many immigrant, Native American, Alaska Native, and Native Hawaiian children in the United States as well as the children of immigrant or conquered peoples in other industrialized countries, such as Australia and those of western Europe (Greenfield, Trumbull, et al., 2006). In this chapter, we describe the relevance of the value systems of individualism and collectivism for learning, development, and education. The articulation of these two value systems led to research demonstrating that the home culture of collectivistic children often opposes the individualistic culture of schools (Greenfield, Raeff, & Quiroz, 2000; Raeff, Greenfield, & Quiroz, 2000).

This situation creates the need for educational intervention. The Bridging Cultures Project®, described in the second part of this chapter, is just such an intervention. It was designed to alleviate the cross-cultural value conflict experienced by most immigrant families from Mexico and Central America when they send their children to school in the United States.

On the whole, individualism emphasizes individual success and collectivism emphasizes the success of the group as a whole (Greenfield, Trumbull, et al., 2006). In individualistic cultures, when asked to describe themselves, people tend to list trait labels referring to aspects of their personalities, such as "hard-working," "intelligent," or "athletic" (Triandis, Brislin, & Hui, 1988), as the teacher in the example above expected. In collectivistic cultures, people are more likely to embed their own personal goals with those of the group, such as their extended family or religious group (Brislin, 1993) and to think of themselves as defined by their connections to others (Markus & Kitayama, 1991). This is more or less what the boy in the opening example was doing when he tried to talk about other members of his family rather than about himself. Not surprisingly, the United States, according to some measures, is the most individualistic country in the world (Hofstede, 2001). However, this developmental pathway is hardly universal—in the 1980s, some 70% of the world's cultures could be described as collectivistic (Triandis, 1989). While increased wealth, urbanization, technology, and formal education have driven cultures and individuals around the world in an individualistic direction since the 1980s (Greenfield, 2009), major cross-cultural differences still exist.

ETHNOTHEORIES AND DEVELOPMENTAL PATHWAYS: AN EXPLANATION FOR CULTURAL VALUES AND SOCIALIZATION PRACTICES

An ethnotheory of development is an implied notion of the ideal child accompanied by beliefs about socialization practices that will produce this ideal (Goodnow, 1988; McGillicuddy-De Lisi & Sigel, 1995). Though psychologists hypothesize, test, and report the *explicit* aspects of child development, parents and teachers (and most everyone else) use *implicit* conceptions based on cultural values that are shared and negotiated among members of cultural communities. Research on ethnotheories has uncovered individualism and collectivism (also referred to as *independence* and *interdependence*) as core dimensions, applicable to many developmental domains (Chao, 1994; Guiterrez & Sameroff, 1990; Yovsi & Keller, 2003). Participants from nonwestern cultural communities, such as Chinese (Chao, 1994), Japanese (Rothbaum, Weisz, Pott, Miyake, & Morelli, 2000), Indian (Saraswathi, 1999; Keller, Voelker, Yovsi, & Shastri, 2005), West African (Ogunnnaike & Houser, 2002, for Nigeria & Nsamenang, 1992; Yovsi, 2001; and Keller et al., 2004, for Cameroon), or Puerto Rican (Harwood, Schoelmerich, Ventura-Cook, Schulze, & Wilson, 1996)—among many others—hold parental ethnotheories that express the cultural ideal of collectivism (Greenfield, Trumbull, et al., 2006). These ethnotheories stress social goals related to loyalty to family, respect for elders, politeness, and responsibility for social and cognitive domains (Harwood, 1992). In Western industrialized cultural groups, such as Germans (Keller et al., 2005), European Americans (Harwood et al., 1996) or the Dutch (Harkness, Super, & van Tijen, 2000), parental ethnotheories express the cultural ideal of individualism. These ethnotheories stress individual achievement and independence along with creativity, curiosity, and assertiveness (Greenfield, Trumbull, et al., 2006).

Socialization practices based on a particular ethnotheory of development begin at birth or even before. For example, in European American culture, expectant parents often prepare a nursery—the baby's own room—to set the stage for the development of the child's independence. This is in contrast to collectivistic cultures that utilize a family bed, where the baby sleeps with the parents.

But ethnicity or national origin alone is not sufficient to predict a person's ethnotheory. As mentioned earlier, influences such as higher socioeconomic status and formal education are associated with a more individualistic orientation (Keller, 2007; Palacios & Moreno, 1996; Tapia Uribe, LeVine & LeVine, 1994). Yet despite these external influences, historical and implicit cultural values persist in identifiable ways across various socioeconomic and educational backgrounds (Keller et al, 2005; Harwood et al., 1996). Ultimately, developmental pathways refer to *relative* group differences. In addition to differences between groups, there are individual differences within every group, especially in complex modern societies. These between- and within-group differences are a matter of emphasis; no one person in a modern society is either completely individualistic or collectivistic.

Table 10.1 describes some major differences between the individualistic and collectivistic pathways of learning and development revealed by research (Greenfield, Trumbull, et al., 2006). We have selected features of each pathway that are most relevant to formal education. These pathways appear and function in ways that differentially value intelligence and knowledge. For example, in more individualistic cultures, cognitive, academic, and scientific knowledge is highly valued, particularly the accumulation of factual knowledge. Independence is demonstrated in school when children work alone, show what they know through speaking out and expressing themselves, and expect praise or other tangible rewards for doing so. Incidentally, the importance of praise and rewards also applies to informal education at home—for example, household chores—as well as to school-based activities. In parallel fashion, the material world is also conceptualized in relation to the individual. Children have individual toys and spaces; sharing takes place by permission of the owner rather than simply being assumed.

Table 10.1

CONTRASTING CULTURAL PATHWAYS OF LEARNING AND DEVELOPMENT

DOMAIN	INDIVIDUALISTIC PATHWAY	COLLECTIVISTIC/ SOCIOCENTRIC PATHWAY
Ethnotheory	Independence, individual success	Interdependence, group/ family success
Valued intelligence	Cognitive/academic/ scientific	Social/relational
Valued knowledge	Physical world, factual knowledge	Social world, narrative knowledge
Models of learning	Independent, active participation, praise	Working in groups, observation, criticism
Communication	Speaking, self-expression	Comprehending, speech that is respectful to authority
Material world	Personal ownership, sharing by choice	Shared use, responsibility to share

Adapted from "Cultural Conceptions of Learning and Development," by P. M. Greenfield, E. Trumbull, H. Keller, C. Rothstein-Fisch, L. Suzuki, & B. Quiroz, 2006. In P. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology* (2nd ed., pp. 675–692). Mahwah, NJ: Erlbaum.

In contrast, collectivistic cultures value social intelligence as it relates to people, not facts or things: It is situated in a social world where knowledge about people's experiences is highly valued. Children are socialized to become interdependent with others. They work together to help and share with other members of the group, instead of being showcased for their individual achievement (Rothstein-Fisch & Trumbull, 2008). Praise may also make them feel singled out and uncomfortable rather than make them feel good about themselves (Greenfield, Quiroz, & Raeff, 2000).

FROM THEORY TO RESEARCH: APPLYING THE TWO CULTURAL PATHWAYS OF DEVELOPMENT TO FORMAL EDUCATION

Two empirical studies document cross-cultural value conflict between the cultural pathway of development assumed by Latino immigrant families and that taken for granted by the schools. In both studies, immigrant parents demonstrated a much stronger collectivistic orientation to child socialization than their children's teachers (Greenfield et al., 2000; Raeff, et al., 2000). For example, Greenfield and colleagues (2000) found that immigrant Latino parents of third and fourth-grade students and their European American teacher were often confused by each other's comments during naturally occurring parent–teacher conferences. Most conversations between the parents and the teacher did *not* confirm or elaborate a common theme—a phenomenon referred to as "noncooperative discourse." Instead, there were frequent instances where parents and teachers seemed to completely miss what the other was trying to say. More important, the implicit values related to expectations for learning, development, and behavior that underpinned each participant's approach to the topics under discussion seemed to be unknown to the other. These differences between the teacher's values and those of the parents coalesced into five major themes. In each case, the teacher's orientation (listed first) represented individualistic values whereas the Latino immigrant parents' orientation (cited second) reflected collectivistic values. These differences included (a) individual versus family accomplishment; (b) praise versus criticism; (c) cognitive versus social skills; (d) oral expression versus respectful communication with authority; and (e) parents' role in teaching the child versus the parents' role in socializing the child. Overall, this study demonstrated that cultural values and the behaviors that emanate from them were not shared between the parents' and their child's teacher.

In the second study, fifth-grade children, their parents, and their teachers were given social dilemmas based on the real-world experience of immigrant Latino families in Los Angeles (Raeff et al., 2000). These dilemmas were likely to be resolved in either an individualistic or collectivistic manner. In the School One sample, the children and parents were all European American, the teachers were predominantly European American. In this setting, the scenarios were solved by children, parents, and teachers in a predominantly individualistic manner. This was not the case with School Two, where all of the families were Latino immigrants, again the teachers were predominantly European American. Data from responses to several scenarios showed that the Latino parents were more concerned about their children's sharing and helping, while the teachers expressed a greater orientation toward task completion (get the job done), individual choice (the student can chose to help or not help), and personal property (maintaining personal ownership versus sharing).The data from this study demonstrate that, as in the study of parent-teacher conferences, children are pulled in two directions: having to make choices about their behavior that either coincide with the expectations of their families or that align with those of their teachers.

Note that in both schools there was a minority of teachers who were not European American (including two Latino teachers in each school). However, a subsequent study (described in more detail below) confirmed that teachers generally give individualistic responses no matter what their ethnic background (Trumbull, Rothstein-Fisch, Greenfield, & Quiroz, 2001). Our conclusion concerning teachers' values is that formal education and especially teacher training exert a strong pull toward individualistic values. Indeed, the average educational level of the Latino immigrant parents was fifth grade, obviously a much lower level of formal education than that possessed by Latino teachers in the two schools (Trumbull et al., 2001). In contrast, the European American parents had, on the average, completed 4 years of postsecondary education, very comparable to the educational level that would be expected of teachers. The role of formal education as an individualizing influence was confirmed by a later study comparing the ethnotheories of Latino immigrant nannies and their employers in Los Angeles. There, it was very clear that formal education trumped ethnicity or national origin as a primary influencing factor on cultural beliefs and practices relating to early care and informal education (Greenfield, Flores, Davis, & Salimkhan, 2008). Thus, as the level of formal education increases around the world, we can expect increasing individualism in parental ethnotheories (Greenfield, 2009).

FROM RESEARCH TO INTERVENTION: THE *BRIDGING CULTURES PROJECT*

Based on ethnographic observation, the study of parent–teacher conferences, and the scenario study, four researchers (Patricia M. Greenfield, Blanca Quiroz, Carrie Rothstein-Fisch, and Elise Trumbull) posed the question: How might knowledge of the cultural pathways of individualism and collectivism impact teachers' ability to work with students and families with a collectivistic values system? Because our research was based on Latino immigrant families and their school experience, we selected schools serving predominantly Latino immigrant students and families as the target of our intervention. We reasoned that teachers with large numbers of immigrant Latino students might be best situated to see overt cultural differences between the collectivistic immigrant Latino families and the individualistic orientation of schools. Therefore, we recruited seven bilingual (Spanish-English) elementary school teachers from the greater Los Angeles area (see Trumbull, Diaz-Meza, Hasan, & Rothstein-Fisch, 2001). Four of the teachers were Latino and three were European American; six were female, one was male; all grade levels of elementary school were represented, kindergarten through fifth grade. Teachers' years of teaching experience ranged from 5 to 21 years (average 12.7 years).

Methods of the *Bridging Cultures Project*

The seven teachers and four researchers participated in three half-day workshops approximately one month apart. In the first workshop, the teachers completed a pretest using the same individualism–collectivism dilemmas described above (Raeff et al., 2000). The teachers responded with a strongly (86%) individualistic perspective (Trumbull et al., 2001). In the debriefing, the teachers were stunned to learn two things. First, they found out that Latino parents solved the dilemmas differently from the largely European American teachers. Second, the Latino teachers were particularly surprised to see how individualistic they had become: Their responses seemed to align much more with the mainstream values of individualism, at least on the surface, than with the collectivistic values of the students and families (and in most cases their own home culture). As they told us later, they attributed this shift to their own schooling process where the school way was right and the family way was wrong, causing untold confusion and stress. These Latino teachers had sensed the tension but never understood it or given it a name up to that point.

The remainder of the first workshop consisted of more examples of cross-cultural conflict, including those regarding individual versus shared class materials (Quiroz & Greenfield, 1996) and what counts as "valued" knowledge in the classroom (Greenfield, Raeff, & Quiroz, 1996). Between the first and second workshops, teachers were asked to observe examples of individualism and collectivism in their classrooms and school. They returned a month later for the second workshop with many examples highlighting the implicit values of the school. In the workshop, the researchers clarified and elaborated on the teachers' observations. At the end of the second workshop, the teachers were asked to make a change in their classroom practice with the goal of reducing a conflict between individualistic and collectivistic practices. One month later, the teachers came back to the group meeting abuzz. They reported that they had made changes in classroom management and in their relations with families. This marked the beginning of a process through which recognition of the individualism–collectivism pathways could generate new classroom practices to facilitate learning. It is an important feature of the Bridging Cultures teacher development process that the researcher/trainers provide the theoretical paradigm and supporting data, but the changes in practices come from the teachers themselves (Rothstein-Fisch & Trumbull, 2008).

At the end of the third workshop, teachers took a posttest drawn from the same set of dilemmas as the pretest (Raeff et al., 2000). The posttest revealed that the teachers had made a significant shift in thinking about how to resolve home and school conflict: Collectivistic responses increased from 14% to 57%, while individualistic responses dropped from 86% to 21%. Likewise, 21% of the responses in the posttest combined individualistic and collectivistic elements. In addition to representing many more collectivistic responses, teachers' solutions were much better distributed between the two value systems. The teachers initiated the idea and unanimously agreed to continue meeting (as was our hope), with one teacher's comment speaking for the whole: "Meeting three times sets the fire, but nothing's been cooked yet. The risks are that people will go back and close the door on their classrooms. We should keep this core group alive" (Trumbull, Diaz-Meza, Hasan, & Rothstein-Fisch, 2000, p. 13). Over the next 5 years, the researchers met with the teachers every 2 or 3 months. The meetings always included at least one meal because we valued the social as well as the intellectual components of the relationships within the group. In this way, we were modeling the integration of social and the cognitive values that was necessary to bridge the two cultural paradigms (see Table 10.1).

In addition to the whole-group meetings, we gathered information through classroom observations of the seven teachers at least twice, including a debriefing after each visit. We also conducted detailed interviews with all seven teachers, gathering insights into their teaching innovations, the effects of their innovations on students' learning, and reflections on their own teaching practices. These methods (direct classroom observations, lengthy, reflective individual interviews, and meetings of the researchers and teachers) yielded an extensive body of data for all seven teachers.

Teachers as Researchers

Almost immediately, the teachers became teacher–researchers themselves. They experimented with new ways of using their knowledge of individualism and collectivism to modify and adapt their practices. During group meetings they reported their experiences for others, thus keeping their own thinking fresh and helping others to learn from their experiences. We believe that teacher research is an important and unique source of knowledge about teaching and learning (Cochran-Smith & Lytle, 1999; Dinkelman, 1997). It follows that artificial boundaries between the practice of teaching and research on teaching need to be challenged. Much can be gained from collaborations between researchers, educational psychologists, professional development specialists, and classroom teachers.

The Bridging Cultures teachers proved to be important partners in the dissemination process. Our seven teachers have disseminated locally (in their schools and at the district level), regionally, and nationally, as have the researchers. The teachers have also contributed to the broader educational community through collaborative publications with the researchers (cf. Quiroz, Greenfield, & Altchech, 1998, 1999; Rothstein-Fisch, Trumbull, Isaac, Daley, & Pérez, 2003; Trumbull, Rothstein-Fisch, & Hernandez, 2003).

Ethnography: A Useful Method for Understanding Families

Another aspect of teacher research that was part of the Bridging Cultures professional development process was the use of ethnography, a technique from anthropology that is often defined as participant observation (Trumbull et al., 2001). We encouraged teachers to get to know their students' parents and to learn more about their backgrounds—including how much access parents had had to formal education in their home countries. The teachers learned that most of their students' parents had little opportunity to go beyond sixth grade in Mexico or Central America, and this limited formal education then became a barrier to helping their children with homework and the development of academic skills. Individual family differences—as a function of acculturation level, economic level, and educational level—were revealed through ethnography, while the individualism–collectivism paradigm provided a framework through which the developmental goals that motivated observed differences could be understood.

FROM INTERVENTION TO INNOVATION: HOW *BRIDGING CULTURES* CHANGED CLASSROOM PRACTICES

The teachers' experimentation in their own classrooms and schools has shown the framework of individualism and collectivism to be more generative than we ever dreamed possible, resulting in vast numbers of innovations to support student learning. We have identified many of these classroom practices elsewhere (e.g., Rothstein-Fisch, 2003; Rothstein-Fisch & Trumbull, 2008; Trumbull et al., 2003). The innovations highlighted in this chapter pertain to math, language arts, science, classroom management, and relations with parents.

Utilizing Group Motivation to Achieve Learning in Mathematics

During one classroom visit, one of the core Bridging Cultures researchers noticed an interesting classroom chart. It was a graph displaying children's names and corresponding stars next to each name indicating the level of memorized multiplication facts they had mastered. This seemed like a curiously individualistic motivational tool for children of migrant farm workers. When asked about the star chart during the observation debriefing interview, the teacher explained that the students were doing poorly in learning timed multiplication facts. She was having difficulty motivating the students, and she fell back on the star chart idea she had learned during teacher training. At the time she introduced the star chart, she conducted a class meeting that allowed her to share her concern for students' lack of math progress. She gave the students an opportunity to look at the chart and talk about it. The teacher reports,

> The students said, "Wouldn't it be neat if it would be a solid block of stars, and the whole chart was filled in," and everybody said "Yeah, yeah, that would be so neat." The students started to say they wanted to help each other. Everyone who needed help got adopted by students who had already mastered [the work]. They started helping each other pass, and they seemed to move ahead. The [more advanced] buddies put their own learning on hold in order to help their [less advanced] buddies, not for individual success, but for the success of the group. (Rothstein-Fisch et al., 2003, p. 132)

Thus the children were motivated to help each other study timed math facts (so that the whole group could achieve proficiency) much more than to work on increasing their own individual accumulation of stars, indicating their individual competence (Rothstein-Fisch & Trumbull, 2008).

Their motivation was rooted in the value of group success rather than in objects or individual awards. Incidentally, when a student was ready to be tested by the teacher at the back of the room, the teacher allowed the buddy to observe, not to provide answers but just to show support and provide encouragement. If the student was not successful, the buddy knew where to help out during the next study session. If the student was successful, the buddy would ring a special bell. This signaled to the entire class that another star had been added to "their" chart, and the children all stopped their work to applaud the individual's contribution to the group's success.

We observed this process several times in the classroom: the bell ringing to mark a shared success never seemed to interrupt the classroom but rather it alerted the class to the progress being made for the collective good. According to the teacher "In third grade, they only have to go up to the 5s [tables], but many went to the 12s! All got to the 6s . . . they went beyond the requirement (Rothstein-Fisch et al., 2003, p. 133).

This example highlights how students were motivated to work for the good of the group rather than for their own individual success. Moreover, it appears that they had never conceptualized the star chart as meaningful as an individual enterprise in the first place. Filling in the chart with stars made sense to them only when they understood it as something for the whole group to achieve. But without an understanding of collectivism (or, as the teacher in this example likes to say, "the power of the group"), the teacher would have been unlikely to allow the students to construct their own meaning of the chart. Likewise, it is highly improbable that the students could have accomplished so much math success on their own. Most important, this example demonstrates that academic goals can be achieved (or in this case exceeded) if children's home cultures are understood and treated as sources of learning strengths.

Another example of honoring the children's collectivistic home culture in the service of mathematics learning occurred in an activity designed to teach the meaning of the concept 1,000. In one class, children were organized into table groups of four, each responsible for collecting 1,000 of the same kind of object of their own choosing. Children volunteered to help not only their own group but all the groups, saying things like, "I can bring in pennies to help Table 3"; "I can bring in rocks"; "I can bring in plastic bags." When the teacher asked them what they were going to do next, the third-grade children said, "Get our stuff out, help each other, and count" (Rothstein-Fisch & Trumbull, 2008, p. 134).

In mainstream classrooms, it is likely that teachers would set up collections by individual students or perhaps teams of four children, but these might be established as a competition—to see which person or table could get 1,000 objects collected first. This kind of teacher might be reasoning that an individual prize would be motivating for the students. However, in the case of these immigrant Latino children, it appeared that the "helping of others" was the true incentive. Teachers who think that they need to provide rewards for success in a competition may be undermining students' collaborative goal, a goal that ultimately helps children from more collectivistic home backgrounds to learn. Whereas a teacher might think that offering an incentive would speed up the collection of items and thus the opportunity to learn about the concept of 1,000, just the opposite is likely to be true for children raised with a more collectivistic value system. Because all the children helped their classmates as a whole, everyone's learning was likely to be enhanced.

This point is evidenced in the children's discussion of their collection and counting processes:

> The groups reported on how many objects they had gathered and how many more they needed. They also described the method they used to count their objects. This activity led other students to know how many more pennies, rocks or nails to gather (identifying which group needed the most help), while allowing the children to see the many different ways that objects could be grouped. Discussion centered on why certain objects were easier to count in larger or smaller groups, making explicit the strategies behind the choices the children had made. (Rothstein-Fisch & Trumbull, 2008, p. 134).

Cultural Knowledge Facilitates Learning in Language Arts and Science

Bridging Cultures teachers were also successful in promoting language development, which was particularly important because many of their students were English learners. Frequent use of choral reading activities allowed students to practice their burgeoning English skills without fear of error because their voices could blend in with the group (Rothstein-Fisch & Trumbull, 2008). Oral language development is an important academic goal. However, children from collectivistic families are generally socialized by their parents to show respect by listening and not speaking up. The teachers reported that their students seemed uncomfortable even when discussing an object brought from home on sharing day. To mitigate students' discomfort, teachers allowed students to come to the front of the room in small groups of four to discuss their special object. Thus, meeting the individualistic goal of learning to "speak out" was facilitated by the teachers' application of knowledge that group support, at least initially, would help students.

Another aspect of promoting language development centered around the children's need to tell stories. Narratives, especially narratives involving family, are a valued form of discourse in Latino immigrant culture. The value placed on narrative discourse contrasts with the almost exclusively cognitive or academic discourse emphasized in school—often presented in the form of decontextualized facts, skills instruction, or topic-restricted discussions (see Gee, 1989; Heath, 1983; and Michaels, 1981, for discussions of different narrative styles). All seven teachers found ways to incorporate children's stories into their instruction. In one classroom, children shared stories with seat mates as a way of tapping prior knowledge about a particular topic before the teacher presented the lesson in math, science, or social studies. In this way, everyone had a chance to share a story, but it required only a fraction of the class time compared to what would be required if everyone in the whole class took a turn. As the stories were being told to seat mates, the teacher circulated around to hear snippets of each story allowing him to adapt the transition to "academic" knowledge in his lesson by pointing out a few key points from various stories.

In one fourth-grade classroom, the children seemed particularly eager to discuss their grandmothers, many of whom still lived in Mexico or El Salvador. To honor their interest and to facilitate language development, the teacher helped the children to craft a series of questions they could all answer in describing their grandmothers: where their grandmother lived, how old she was, what foods she liked to cook, what they liked to do with their grandmother, and how much they loved their grandmother. These papers were drafted, edited, and prominently displayed on the classroom bulletin board. In this way, children's identified interests were tapped; they learned how to write and edit a descriptive paragraph; and their work was shared for everyone to read (Rothstein-Fisch & Trumbull, 2008).

Another example illustrates how students are able to connect their narrative knowledge with scientific knowledge. The children in fifth grade were studying the wetland habitat and preparing for a field trip to see it up close. A docent from the Audubon Society was visiting the classroom to help the children learn more about the birds of the wetlands and he asked the children what they knew about birds. The students responded with stories about their family experiences, including trips to the park or beach. After a few family stories, the frustrated visitor screeched: "No more stories!" and the children's voices fell silent.

After the docent left, the teacher knew exactly what she wanted to do. Based on a discussion from a Bridging Cultures group meeting, the teacher recalled the "egg" example wherein the teacher asked the children "what eggs look like when they are cracked." One child tried several times to tell about a time she had cooked and eaten eggs with her grandmother, but the child was overlooked in favor of a child who described eggs as white and yellow (Greenfield et al., 1996). With the egg example in mind, the Bridging Cultures teacher drew a large

"T" chart on the board. She began with the key phrases from the children's stories on the left side of the chart. One student commented that while she was in the garden with her grandmother, she noticed a hummingbird flying yet staying in one place. This observation produced a fascinating discussion and led to the facts, placed in the right-hand column, that hummingbirds can "hover" and fly in any direction. To fly in one place (i.e., hover) required the hummingbird's wings to beat very rapidly, demanding large amounts of energy. How do the hummingbirds sustain this kind of energy? The children reasoned that hummingbirds must have to eat frequently and the topic of metabolic rates ensued (Rothstein-Fisch, 2003). Because the learning was embedded in the social world of a classmate, it was likely to increase students' interest in the subject and, arguably, to improve learning (Alexander & Jetton, 2003).

Classroom Management

Changes in classroom management seemed to be affected early on and most dramatically. Like approaches to learning, approaches to classroom management are culturally situated. The cultures of the school, teacher, and the students are all systems at work affecting every kind of classroom management issue, from establishing classroom rules (Rothstein-Fisch & Trumbull, 2008) to deciding what constitutes helping versus cheating (Rothstein-Fisch et al., 2003). In some cases, lengthy classroom rules were shortened to emphasize the overarching goal of "respecting others." The value of respect (*respeto)* has rich meaning for children from Latino immigrant families: Respecting others is central to their socialization (Valdés, 1996). Thus when these children list "respect others" in their classroom rules, the phrase is likely to conjure up a wide range of appropriate behaviors, such as listening to the teacher, complying with the teacher's requests, sharing materials, and not offending another person's sense of dignity.

Room arrangement and the physical aspects of the classroom also reflected a collectivistic perspective in the Bridging Cultures classrooms. In the lower elementary grades, children sat very close to each other during rug time, not on individual carpet squares as is often the case in classroom practice. The children were observed touching each other's hair or shoes in a completely nondisruptive way, much as they might do in their homes with siblings or cousins. Because the teachers now understood this to be natural they did not have to take away from lesson time to say, "Keep your hands to yourself," instead they let the children behave as they would naturally—allowing them to stay focused rather than separated and perhaps less comfortable (Rothstein-Fisch & Trumbull, 2008).

Changed Practices With Families

In the United States, the No-Child-Left-Behind legislation indicates that "parents are seen as the child's first and most important teachers, and for children to be successful in school, parents and families need to be actively involved in their children's learning" (U.S. Department of Education, 2007). However, as the study of parent–teacher conferences demonstrated, there was genuine disconnection (in 21 of the 22 times this topic came up) between what the parents thought their role was and what the teacher thought they should do (Greenfield et al., 2000). According to the researchers, "Although the parents did not want to teach at home, they did want to maintain their jurisdiction as socializing agents at home" (p. 105).

Understanding that parents might not be able or inclined to "teach" their children at home, one teacher allowed children to practice and discuss homework in small groups without writing anything down (Rothstein-Fisch & Trumbull, 2008). The children were observed to discuss the homework questions (in this case about a story from their language arts textbook) and decide what would be the best answer for each question. This allowed the children more time to think about and reflect on what they had read, discuss the questions in the group, and essentially practice the homework, thereby making it far easier to complete the homework without any assistance at home.

Bridging Cultures teachers almost immediately initiated changes in how they communicated with and supported families (Trumbull et al., 2003). One teacher divided her students' families into three groups for the conference: two for Spanish-speaking parents and one for English-speaking parents. According to the teacher,

> a warm feeling came across during the conferencing. Many parents had questions that benefited the others. . . The group conferencing was relaxing for the parents. It was a less threatening environment than the individual style; parents supplied support and were company for one another. This format provided a group voice for the parents rather an individual voice. After one hour, parents could sign up for a private conference or ask a few questions privately. (Quiroz et al., 1999, p. 69)

As a result of this teacher's success, her principal asked her to conduct a workshop at her school to encourage other teachers to try the group conference approach. Other teachers have picked up on this practice and continue to experiment and adapt it to their own settings.

In another case, one teacher decided she would try to increase the number of her parent volunteers. Although the teacher is Latina and her first language is Spanish, she reports,

> Both the parents and I had difficulty approaching each other for help. Most parents had little formal education and probably did not know they could actually assist in the classroom: only a few had attended junior high or high school. I had to conduct my own informal ethnographic research about my families and began to build relationships with parents in the process. . . . Although I was now averaging five parent volunteers a week, I still felt like there was something missing. Many parents would stay but were uncomfortable [interrupting me] while I was teaching a lesson and ask what they could do. (Trumbull et al., 2001, pp. 85–86)

To support parents, the teacher compiled a notebook with a paragraph describing why parental help would be so useful in the class. She also described an array of different ways parents could help in the classroom, from cutting out materials to help with reading. Younger siblings also were welcome in the class because this would allow more parents to participate while providing younger children with some exposure to elementary school.

We observed the parent volunteers in action during our visits to her classroom, each checking in and out via the parent notebook. On one day, there were four parent volunteers. One was helping a group of children read, while another was watching and learning from that parent how to support this kind of learning. In this way, the teacher set up a scaffold for classroom volunteers: The more experienced volunteers showed the newer ones how to use materials and support student learning. This arrangement had multiple benefits. First, because of the volunteers, the children were getting more help (not incidentally, something the parents were likely to value greatly, both in being the person able to help and in knowing the children were getting assistance). Second, parents otherwise unfamiliar with the U.S. system of education gained an intimate, up-close view of how a classroom was set up and organized, they also became familiar with the teacher's expectations for learning. Third, because the younger siblings were also included in the group—with the kindergarten, first-, and second-grade students—they too were able to preview what school would be like for them while at the same time being able to participate as part of the family group. Ultimately, the teacher more than doubled her volunteers from 5 to 12, 10 of whom put in over 100 hours during the school year (Trumbull et al., 2003).

BEYOND THE ORIGINAL *BRIDGING CULTURES* TEACHERS

The Bridging Cultures paradigm, based on the concepts of individualism and collectivism, has been applied to other educational audiences beyond the seven original teachers. Our team has carried out professional development with preservice teachers (Rothstein-Fisch, 2003) and early childhood educators (Zepeda et al., 2006), with participants indicating that they found the framework applicable to their work with typical and special needs populations (Rothstein-Fisch, 2006, 2007a,b,c). School counselors have also found the framework useful (Geary, 2001; Roman, 2006). Current research is underway to explore the impact of the individualism-collectivism framework on students studying elementary, secondary, and special education, educational psychology and counseling; and educational leadership and policy studies at California State University, Northridge.

The innovations developed by teachers are seemingly endless. Teachers can apply the framework of individualism and collectivism in ways that make sense in their classrooms and schools and with which they are comfortable. Not all innovations have been equally valuable or successful. One linchpin to the success of innovations we describe is the support of school administrators. As teachers create new classroom practices, principals and coworkers (such as paraprofessionals and school specialists) need to support these efforts and not impose unnecessary obstacles. For instance, it was important that the principal allow younger siblings in the classroom in the parent-volunteer example described earlier.

There is no recommended ratio of individualism to collectivism in the classroom, although most of the innovations have, quite naturally, been in the direction of making uniformly individualistic classrooms more collectivistic. However, it is equally important to note that the Bridging Cultures teachers have not rejected individualism; they are just more mindful of when individualistic practices support learning and when they hinder it. From the beginning, they were very aware that eventually their students would have to learn how to succeed in an individualistic educational world. For them, the bridge to individualism was as important as the introduction of more collectivistic practices. Ultimately, because it affords a deeper understanding of culture, the Bridging Cultures framework has allowed teachers to become creative and mindful problem solvers as they try to support learning in ways that do not undercut their students' home values.

It is important to note that our method is nonprescriptive. We provided the paradigm and the teachers used the paradigm to generate their own ideas varying greatly not only from teacher to teacher but also from setting to setting. When the teacher whose parent volunteer activities were described above moved from one school to another, she found that the new social context demanded new strategies for parent involvement. Thus, innovations can be taken up at one time with one group of students and families and jettisoned for another strategy as appropriate with a new group of students and families.

Reasons for Optimism

The outcomes of the Bridging Cultures Project are causes for optimism. Some of the most striking effects have to do with (a) the perspective teachers have gained on their own culture, their students' home culture, and school culture; (b) the degree to which this new perspective has begun to influence their thinking and their practice in ways that reduce conflicts between home and school culture; and (c) the increased confidence teachers have in their own abilities to build the kinds of relationships with families that will support student success in school and family unity at home.

Note that it is highly unusual for professional development efforts to have a documented long-term impact on teacher practice beyond the training itself. We have documented an enduring transformation in the Bridging Cultures teachers in their ability to reach their students and work with families successfully. These teachers know how to learn from their students' families, and they have new ways of understanding what parents are sharing with them. We believe that the project has been successful for the following reasons:

- It uses a theory- and research-based framework to guide practical and applied experimentation with new educational methods.
- The theoretical framework is economical—that is, it has only two organizing concepts, individualism and collectivism. This sim-

plicity makes it accessible. Yet it is simple without being simple-minded or simplistic. Because it is accessible, teachers can use it immediately to guide their work, rather than thinking that culture is entirely too complex and thereby throwing up their hands as if to say "there are so many cultures, I have to just treat each student as an individual." When this happens, educators have essentially defaulted to the mainstream value of individualism and, as a result, have eliminated the role of culture altogether.

- The individualism–collectivism framework has been productively applied to some of the most important and difficult aspects of teachers work, including classroom management, subject matter instruction, assessment, and parent-school relations
- Longitudinal professional development offers teachers opportunities to share and analyze practice over an extended period of time, valuing their knowledge and experience as a group. One cannot know to what degree shorter-term professional development activities would have affected teacher practice and student learning.
- The Bridging Cultures Project included meetings that incorporated both rigorous intellectual work and enjoyable interpersonal activities such as sharing meals, humor, and personal celebrations. This format provided an opportunity (not available in shorter-term interventions) to integrate the intellectual with the social, thus enacting a bridge between the two cultures in the professional development activities themselves.

In the final analysis, teachers recognize that neither value system is all good or all bad. One teacher said, "I think that it is a good point to bring out about culture . . . that . . . we're not saying collectivism is right and individualism is wrong. We're just saying to recognize it. It's different" (Greenfield, Trumbull, et al., 2006, p. 686).

CONCLUSIONS AND FUTURE DIRECTIONS

Culture is not simply a variable to consider in describing learners. It is at the core of decisions made by teachers, parents, school officials, and administrators. As systems of values and beliefs, ethnotheories drive learning, socialization, and development (Greenfield, Suzuki, & Rothstein-Fisch, 2006). The framework of individualism and collectivism allows teachers to understand the values and practices of families that do not share the school's cultural assumptions. It has proven to be a helpful way to begin to rethink formal education in a multicultural context.

Thus far, we have explored the use of intensive qualitative longitudinal methods to document and assess the impact on teachers. We have also begun to explore the impact on parents (Esau, Greenfield, Daley, Robles, & Tynes, 2010). The next step is to explore the impact on students more explicitly, beyond what teachers have reported informally. An important future direction for research would be a longitudinal intervention in which children from immigrant Latino families in one school are exposed to the Bridging Cultures approach for a period of years, from kindergarten through secondary education, while children from similar backgrounds in another comparable school are not. Academic outcomes, degree of comfort in the school climate, and family-school relationships would ideally all be compared across the two groups. This type of extended intervention and controlled research design have great potential to confirm and extend our qualitative findings. We hypothesize that this type of research will demonstrate that the Bridging Cultures approach reduces the conflict children feel between home values and school values, enhances their motivation to learn, and enables them to function harmoniously at home while achieving academically at school.

ACKNOWLEDGMENTS

The Bridging Cultures Project was initially supported by a grant from the U.S. Department of Education, Office of Educational Research and Improvement to WestEd, the regional educational laboratory for the Western United States. WestEd holds the Bridging Cultures® trademark and has licensed its use to the four core Bridging Cultures researchers: Patricia M. Greenfield, Blanca Quiroz, Carrie Rothstein-Fisch, and Elise Trumbull. Additional support has been provided by the Michael D. Eisner College of Education at California State University, Northridge. Portions of this chapter appeared previously in "Cultural Conceptions of Learning and Development," by Greenfield, Trumbull, Keller, Rothstein-Fisch, Suzuki, and Quiroz, 2006. In P. A .Alexander and P. H. Winne (Ed.), *Handbook of Educational Psychology* (2nd ed., pp. 675–692). Mahwah, NJ: Erlbaum.

REFERENCES

Alexander, P. A., & Jetton, T. L. (2003). Learning from traditional and alternative texts: New conceptualizations for the information age. In A. L. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 199–242). Mahwah, NJ: Erlbaum.

Brislin, R. (1993). Understanding culture's influence on behavior. Fort Worth, TX: Harcourt Brace.

Chao, R. K. (1994). Beyond parental control and authoritarian parenting style: Understanding Chinese parenting through the cultural notion of training. *Child Development, 65,* 111–119.

Cochran-Smith, M., & Lytle, S. L. (1999). The teacher research movement: A decade later. *Educational Researcher, 28*(11), 15–25.

Dinkelman, T. (1997). The promise of action research for critically reflective teacher education. *Teacher Educator, 32*(4), 250–274.

Esau, P., Greenfield, P. M., Daley, C., Robles, F., & Tynes, B. (2010). Developing cross-cultural harmony in schools serving Latino immigrant families in Los Angeles: Bridging Cultures parent workshops. In G. Marsico, K. Komatsu, & A. Iannaccone (Eds.), *Crossing boundaries: Intercontextual dynamics between family and school.* Manuscript in preparation.

Geary, P. (2001) *Bridging cultures with middle school counselors.* Unpublished master's thesis. California State University, Northridge.

Gee, J. P. (1989). Two styles of narrative construction and their linguistic and educational implications. *Discourse Processes, 12,* 287–307.

Goodnow, J. J. (1988). Parents' ideas, actions, and feelings: models and methods from developmental and social psychology. *Child Development, 59*(2), 286–320

Greenfield, P. M. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology, 45,* 401–418.

Greenfield, P. M., & Bruner, J. S. (1966). Culture and cognitive growth. *International Journal of Psychology, 1,* 89–107.

Greenfield, P. M. & Cocking, R. (Eds.). (1994). *Cross-cultural roots of minority child development.* Hillsdale, NJ: Erlbaum.

Greenfield, P. M., Flores, A., Davis, H., & Salimkhan, G. (2008). What happens when parents and nannies come from different cultures? Comparing the caregiving belief systems of nannies and their employers. *Journal of Applied Developmental Psychology 29,* 326–336.

Greenfield, P. M., Keller, H., Fuligni, A., & Maynard A. (2003). Cultural pathways through universal development. *Annual Review of Psychology, 54,* 461–490.

Greenfield, P. M., Quiroz, B., & Raeff, C. (2000). Cross-cultural conflict and harmony in the social construction of the child. In S. Harkness, C. Raeff, & C. M. Super (Eds.), *Variability in the social construction of the child* (pp. 93–108). New Directions in Child Development, No. 87. San Francisco: Jossey-Bass.

Greenfield, P. M., Raeff, C., & Quiroz, B. (1996). Cultural values in learning and education. In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices* (pp. 37–55). Alexandria, VA: Association for Supervision and Curriculum Development.

Greenfield, P. M., Suzuki, L., & Rothstein-Fisch (2006). Cultural pathways through human development. In W. Damon & R. M. Lerner (Series Eds.) & I. E. Sigel & K. A. Renninger (Vol. Eds.), *Handbook of child psychology. Vol. 4: Child psychology in practice* (6th ed., pp. 655–699). New York, Wiley.

Greenfield, P. M., Trumbull, E., Keller, H., Rothstein-Fisch, C., Suzuki, L., & Quiroz, B. (2006). Cultural conceptions of learning and development. In P. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 675–692). Mahwah, NJ: Erlbaum.

Guiterrez, J., & Sameroff, A. J. (1990). Determinants of complexity in Mexican-American and Anglo-American mothers' conceptions of child development. *Child Development, 61,* 384–394.

Harkness, S., Super, C., & van Tijen, N. (2000). Individualism and the "western mind" reconsidered: Parents' ethnotheories of the child. In S. Harkness, C. Raeff, & C. M. Super (Eds.), *Variability in the social construction of the child* (pp. 23–39). New Directions in Child Development, No. 87. San Francisco: Jossey-Bass

Harwood, R. L. (1992). The influence of culturally derived values on Anglo and Puerto Rican mothers' perceptions of attachment behavior. *Child Development, 63,* 822–839.

Harwood, R. L., Schoelmerich, A., Ventura-Cook, E., Schulze, P. A,, & Wilson, S. P. (1996). Culture and class influences on Anglo and Puerto Rican mothers' beliefs regarding long-term socialization goals and child behavior. *Child Development, 67,* 2446–2461.

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms.* New York: Cambridge University Press.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.

Keller, H. (1997). Evolutionary approaches. In J. W. Berry, Y. H. Poortinga, & J. Pandey, (Eds.), *Handbook of cross-cultural psychology, Vol. 1: Theory and method* (pp. 215–255). Boston: Allyn & Bacon.

Keller, H. (2003). Socialization for competence: Cultural models of infancy. *Human Development, 46,* 288–311.

Keller, H. (2007). *Cultures of infancy.* Mahwah, NJ: Erlbaum.

Keller, H., Hentschel, E., Yovsi, R. D., Abels, M., Lamm, B., & Haas, V. (2004). The psycho-linguistic embodiment of parental ethnotheories: A new avenue to understand cultural differences in parenting. *Culture & Psychology, 10,* 293–330.

Keller, H., Voelker, S., Yovsi, R. D., & Shastri, J. (2005). The representation of independent and interrelated conceptions of caretaking. In P. Mohite (Ed.), *Theoretical approaches to early development: Implications for intervention.* Baroda, India: University of Baroda.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253.

McGillicuddy-De Lisi, A. V., & Sigel, I. E. (1995). Parental beliefs. In M. H. Bornstein (Ed.), *Handbook of parenting. Vol 3: Status and social conditions of parenting* (pp. 333–358). Hillsdale, NJ: Erlbaum.

Michaels, S. (1981). "Sharing time": Children's narrative styles and differential access to literacy. *Language in Society, 10,* 423–442.

Nsamenang, A. B. (1992). *Human development in cultural context: A third world perspective.* Newbury Park, CA: Sage.

Ogunnaike, O. A., & Houser, R. F. (2002). Yoruba toddlers engagement in errands and cognitive performance on the Yoruba Mental Subscales. *International Journal of Behavioral Development, 26,* 145–153.

Palacios, J., & Moreno, M. C. (1996). Parents' and adolescents' ideas on children: Origins and transmission of intracultural diversity. In S. Harkness & C. M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions and consequences* (pp. 215–253). New York: Guilford.

Quiroz, B., & Greenfield, P. M. (1996). *Cross-cultural value conflict: Removing a barrier to Latino school achievement.* Department of Psychology, UCLA. Unpublished manuscript.

Quiroz, B., Greenfield, P. M., & Altchech, M. (1999) Bridging cultures with a parent-teacher conference. *Educational Leadership, 56,* 68–70.

Raeff, C., Greenfield, P. M., & Quiroz, B. (2000). Conceptualizing interpersonal relationships in the cultural contexts of individualism and collectivism. In S. Harkness, C. Raeff, & C. Super (Eds.), *Variability in the social construction of the child* (pp. 59–74). New Directions in Child Development, No. 87. San Francisco: Jossey-Bass.

Roman, L. (2006). *Bridging cultures: Taking a look at independence and interdependence among high school seniors.* Unpublished master's thesis. California State University, Northridge.

Rothbaum, F., Weisz, J., Pott, M., Miyake, K., & Morelli, G. (2000). Attachment and culture: Security in the United States and Japan. *American Psychologist, 55,* 1093–1104.

Rothstein-Fisch, C. (2003). *Bridging cultures: Teacher education module.* Mahwah, NJ: Erlbaum.

Rothstein-Fisch, C. (June, 2007a). *Training the trainers on cultural influences in early childhood.* Los Angeles: Los Angeles Unified School District.

Rothstein-Fisch, C. (June, 2007b) *Bridging cultures in classroom management: Multiple cultures, multiple voices.* Harrisburg, PA: Pennsylvania Department of Education and the Pennsylvania State University.

Rothstein-Fisch, C., & Albarran, A. (December, 2007c). *Cultural pathways in infant care and development: Why culture matters.* Orlando, FL: Zero to Three National Training Institute.

Rothstein-Fisch, C., & Trumbull, E. (2008). *Managing diverse classrooms: How to build on students' cultural strengths.* Alexandria, VA: Association for Supervision and Curriculum Development.

Rothstein-Fisch, C., Trumbull, E., Isaac, A., Daley, C., & Pérez, A. (2003). When "helping someone else" is the right answer. *Journal of Latinos and Education, 3,* 123–140.

Saraswathi, T. S. (1999). *Culture, socialization, and human development: Theory, research, and applications in the Indian setting.* Thousand Oaks, CA: Sage.

Tapia Uribe, F. M, LeVine, R. A., & LeVine, S. E. (1994). Maternal behavior in a Mexican community: The changing environments of children. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 41–54). Hillsdale, NJ: Erlbaum.

Triandis, H. C. (1989). Cross-cultural studies of individualism and collectivism. *Nebraska Symposium on Motivation, 37,* 43–133.

Triandis, H., Brislin, R., & Hui, C. H. (1988). Cross-cultural training across the individualism-collective divide. *International Journal of Intercultural Relations, 12,* 269–289.

Trumbull, E., Diaz-Meza, R., Hasan, A., & Rothstein-Fisch, C. (2001). *Five-year report of the Bridging Cultures Project.* San Francisco: WestEd. Retrieved from http://www.wested.org/bridgingcultures

Trumbull, E., Rothstein-Fisch, C., Greenfield, P. M., & Quiroz, B. (2001). *Bridging cultures between home and school: A guide for teachers.* Mahwah, NJ: Erlbaum.

Trumbull, E., Rothstein-Fisch, C., & Hernandez, E. (2003). Parent involvement—according to whose values? *School Community Journal, 13,* 45–72.

U.S. Department of Education. *Tools for student success.* Retrieved October 14, 2007, from http://www.ed.gov/parents/academic/help/tools-for-success/tools-for-success

Valdés, G. (1996). *Con respeto.* New York: Teachers College Press.

Whiting, B. B., & Whiting, J. W. M. (1975). *Children of six cultures*. Cambridge, MA: Harvard University Press.

Yovsi, R. D. (2001). *Ethnotheories about breastfeeding and mother-infant interaction. The case of sedentary Nso farmers and nomadic Fulani pastorals with their infants 3–6 months of age in Mbvem subdivision of the Northwest providence of Cameroon, Africa.* PhD thesis. Osnabrück, Germany: University of Osnabrück.

Yovsi, R. D., & Keller, H. (2003). Breastfeeding. An adaptive process. *Ethos, 31,* 147–171.

Zepeda, M., Gonzalez-Mena, J., Rothstein-Fisch, C., & Trumbull, E. (2006). *Bridging cultures in early care and education.* Mahwah, NJ: Erlbaum.

11 Technology-Supported Face-to-Face Small-Group Collaborative Formative Assessment and Its Integration in the Classroom

MIGUEL NUSSBAUM, FLORENCIA GOMEZ, JAVIERA MENA, PATRICIA IMBARACK, ALEX TORRES, MARCOS SINGER, AND MARÍA ELENA MORA

The integration of technologies into school learning processes is motivated not only by their central role in the worlds of work and knowledge but also by their potential as mediators of social relations. They bring powerful tools to the task of converting classroom experiences into interactive and collaborative ones that deliver a range of pedagogical benefits (Wood & Malley, 1996). As Postholm (2007) has stated, the question is not whether information communication technologies (ICTs) can offer the teaching and learning activity but rather how teachers and pupils can approach and use this mediating artifact and benefit from it in their work.

Since ICT resources are found mainly in computer laboratories, activities built around these technologies imply a change in the natural context of classroom teaching and tend to focus on the purely technological aspects (Reynolds, Treharne, & Tripp, 2003). In this sense, the technologies are not truly integrated into the classroom teaching dynamic, and this may limit their impact on teaching styles traditionally used in schools (Watson, 2001).

The role of ICTs in education can be described in terms of the following categories: broadening classroom resources and reference; enhancing working processes and products; mediating subject thinking and learning; fostering more independent pupil activity; and improving pupil motivation to do lessons (Deaney, Ruthven, & Hennessy, 2006). However, the mere incorporation of technologies into the teaching environment is not enough to bring about improvements in education quality (Robertson, 2003). Critics have alleged that technological innovation has not resulted in either real curricular innovation or changes to traditional teaching systems (Soloway et al., 2001) and also maintain that the vast majority of teachers use technology to sustain existing patterns rather than innovate (Conlon & Simpson, 2003; Hayes, 2007). Postholm (2007) exemplifies this last point, noting that even among teachers who agree on the need for integrating constructivist concepts, technology is used to support lecture-based teacher-centered instruction.

It is well established that small-group collaborative activities in which group members work together toward the attainment of common goals are an effective tools for facilitating both academic and social achievement (Dillenbourg, 1999). According to Johnson and Johnson (1989), activities organized to function collaboratively lead to greater achievement and retention than do those with structures that emphasize individual action or competitive behavior.

These findings are grounded in socioconstructivist theories, which hold that learning does not take place in a vacuum but rather within a specific context and through interaction with one's peers (Vygotsky, 1979). Learning is thus understood as a process in which social interaction provides feedback, stimulation, instruction, correction, mutual scaffolding of comprehension, and socially shared construction of meaning (Salomon & Almog, 1998).

However, an effective environment for collaborative learning does not automatically materialize the moment where two or more persons begin working together; for it to emerge, certain conditions must be present that ensure learning is achieved. Adams and Hamm (1996) and Dillenbourg (1999) single out five necessary factors for generating efficient collaborative work: individual responsibility, mutual support, positive interdependence between group members, face-to-face social interaction, and work in small groups.

The role of the teacher in collaborative work is also central, whether it be in the planning of activities or their performance and supervision. The actual carrying out of an activity is the most important and arduous of these tasks (Johnson & Johnson, 1999), given that it involves a change in the conceptualization of the very role instructors adopt in the classroom to one revolving around the monitoring of student learning. This new task is thus centered on the students.

When collaborative work is supported technologically, it is known as computer-supported collaborative learning (CSCL), an approach oriented toward the development of computer programs that facilitate interaction between peers and group work. In a CSCL context the technology mediates the interaction between the participants by delivering information, regulating the tasks to be performed, administering rules and roles and mediating the acquisition of new knowledge (Kumar, 1996). The objective is for the technology to offer a medium for classroom discussions that can facilitate participation and social interaction among the students and between them and the teacher (Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003) while also increasing the effectiveness of interaction among peers (Dillenbourg, 1999).

Generally, collaborative activities suffer from certain limitations stemming from the use of desktop computers (Nussbaum et al., 2007), whose software programs are not designed for such applications and whose capacity is insufficient to support simultaneous interactions between various users (Inkpen, 1999). Furthermore, the sort of collaboration supported by most CSCL applications requires that students gather around a single computer and take turns using the mouse or keyboard (Zurita & Nussbaum, 2004a), thus sticking closely to the traditional "one computer–one person" paradigm (Inkpen, 1999). Working at a desktop machine has the added disadvantage of hindering the face-to-face work that is essential for enhancing interaction in a collaborative activity (Zurita & Nussbaum, 2007).

In light of the above, the use of portable technologies that offer individual computer access (1:1) can be a major source of support for the development of collaborative dynamics given that such devices, if used together with appropriate pedagogical designs, facilitate communication between peers and motivate interactions (Roschelle, Rosas, & Nussbaum, 2005). The growth of these technologies has led to the emergence in the educational technology field of the concept of mobile learning. This learning model, based on mobile computers that support wireless communication, offers a number of undeniable advantages for overcoming the above-described limitations of using desktop machines. These advantages include the following:

- Lower cost. This eases the burden of providing an individual computer to each student, improving coverage by reducing the student/computer ratio (Savill-Smith & Kent, 2003).

- Portability. The mobile devices are small, light, and easy to carry; they can therefore be used anytime and anywhere, including in classrooms (Roschelle, Pea, Hoadley, Gordon, & Means, 2001).
- Face-to-face interaction between students within a single group is permitted (Cortez, Nussbaum, Rodríguez, López, & Rosas, 2005).
- Efficient organization of the learning resources used in a given activity is facilitated (Zurita & Nussbaum, 2004b).

These aspects of the mobile model lay the groundwork for a pedagogical proposition that uses technology to support teaching processes based on collaborative dynamics in which each student has access to a portable computer small enough not to impede face-to-face communication—an arrangement that also ensures the mobility necessary to allow random formation of small groups within the classroom.

In this chapter, we introduce and analyze an approach to face-to-face small-group collaborative work mediated by technology that shifts from an instructor-centered arrangement in which the teacher radiates knowledge before a passive class of students to one where the students are active and work collaboratively in small groups while the teacher acts as a mediator. This pedagogical strategy, known as the Eduinnova methodology, was developed at the Pontificia Universidad Católica de Chile over a period of 10 years. It has since been applied with upwards of 20,000 students and 700 teachers and has been successfully integrated in more than 38 schools in Chile, 1 in Argentina, 8 in Brazil, and 3 in England at various educational and socioeconomic levels. The system has also been employed by SRI International in applied research on mathematics learning at three educational institutions in the United States and by the Chilean Ministry of Education for teacher training. The heart of the present work is a study of the Eduinnova project implemented at schools in two Chilean cities, focusing on the analysis of the results obtained.

After discussing the Eduinnova methodology, we next examine the relationship between this methodology and the concept of formative assessment. Then we outline the implementation methodology of the Eduinnova project. The following section describes the study undertaken to analyze the project's results and outcomes. We next set out the results and analyses of the project and a number of aspects of the results and analyses. Finally, we offer a brief conclusion.

TECHNOLOGY-SUPPORTED FACE-TO-FACE SMALL-GROUP COLLABORATIVE LEARNING AND FORMATIVE ASSESSMENT

The Eduinnova methodology for technology-supported face-to-face small-group collaborative learning is inspired by the notion of assessment for learning. When used to adapt teaching to meet learning needs, the concept is known as formative assessment, a process that has been shown to improve students' performance (Black, 2005). It involves the development of learning assessment criteria that enable students to assess the knowledge they are acquiring in order to use it correctly in the real world (Tait, 1997). Two phases can be distinguished in formative assessment (Black & Wiliam, 1998). First, the student perceives a difference between a defined goal and his or her current understanding state; second, he or she takes action to close this knowledge gap and reach the goal.

These two phases are reflected in the functioning of Eduinnova. The formative assessment process begins when a class of students, each supplied with a mobile device, is randomly divided into groups. The teacher, equipped with a specially configured device for monitoring the progress of the entire process, then sends each student a set of multiple-choice questions (MCQs). The group members must individually respond to the questions, thus taking responsibility for doing and assessing their own work. The answers are then presented to the rest of the group, where they are subjected to a peer assessment. The group members attempt to arrive at a consensus answer through discussion, a process facilitated by face-to-face interaction and the small size of the group. All members must contribute and share ideas regardless of what they think of their correctness; the conceptual change must evolve from the learner's preexistent understanding and active involvement in the group discourse (Black, McCormick, James, & Pedder, 2006). If the members do not arrive at a consensus, the system reminds them that they must converge on a single response, in effect forcing them to do so by not permitting them to proceed to the next question. If a group chooses an incorrect response as its consensus answer, the system informs them of their error and instructs them to consider another alternative. Mutual support is the key to this process, as through the collaborative discussion the group members discover where they went wrong, clarify their ideas, and converge upon a new answer based on their individual knowledge and common experience. This loop ends when the group finally selects the right response alternative, at which point they proceed to the next question and repeat the procedure just described until they reach the end of the question set (Cortez et al., 2005). As can be observed in Figure 11.1, group discussion is the nucleus of this activity.

The information obtained by formative assessment is used to provide feedback for modifying the teaching work to meet learning needs (Black, Harrison, Lee, Marshall, & Wiliam, 2004). In Eduinnova this is accomplished through an online in-class management graphic tool (Figure 11.2) incorporated in the teacher's machine that supports his or her mediator role by providing information indicating what the different groups have done well, what must be improved, and how to go about it. The teacher monitors the group outcomes on the tool, to determine which group needs assistance and where they are having difficulties, and can then provide immediate reinforcement or refocus the activity content where required. The screen of the graphic tool shows the groups on the vertical axis and the various questions on the horizontal axis. Each cell formed by the intersection of a group and a question displays one of three different colors, indicating whether an MCQ task was completed correctly on the first attempt, after one mistaken response, or

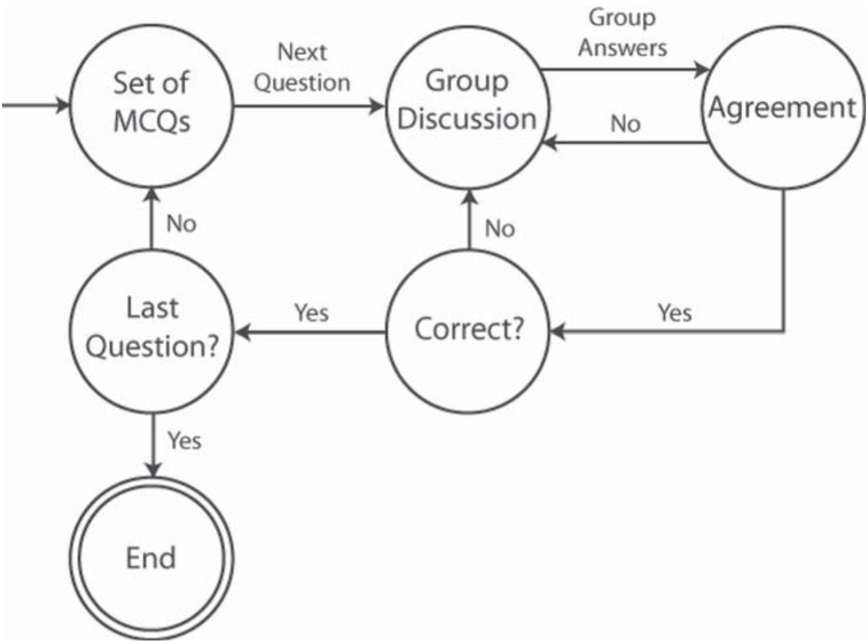


Figure 11.1 Formative assessment dynamics.

after more than one such error. The teacher can also determine whether a given group is developing positive interdependence, suggesting good group work, by observing the speed at which it progresses compared to the other groups. A very rapid advance with (almost) all questions answered correctly might mean that a particularly knowledgeable student has taken control of the group, while very slow headway may be a sign that the group members are not working well together and are therefore having trouble converging on an answer. If the three colors appear in roughly the same numbers, the group may be simply guessing. Thus, the tool allows the teacher to easily determine how well the groups are working, though corroboration by direct observation is still required.

IMPLEMENTATION OF THE EDUINNOVA PROJECT

Project Objectives

The implementation of the Eduinnova methodology in Chilean schools addressed four principal objectives:

- To promote student learning attainment in various curriculum subject areas through a collaborative work methodology mediated by mobile technology.
- To solve the computer coverage problem. Despite major efforts by Chile's Ministry of Education, the computer/student ratio remains low at 1:26 (Enlaces, 2008). The intention is to raise this figure to one notebook per student.
- To integrate the mobile technology into the classroom, the usual daily context of learning. Instead of the class having to go to the technology, the technology leaves its usual work space (labs) and goes to the class.
- To promote interaction between student and teacher through the integration of collaborative activities that deliver performance data to both actors, thus indicating where intervention is required while also strengthening feedback.

Application of the Methodology

The methodology embodies a student-centered approach in which the teacher is seen as an agent promoting the significance and effectiveness of classroom dynamics for learning achievement. Eduinnova works with

Figure 11.2 In-class management graphic tool.

instructors and school authorities to install and implement the pedagogical model and then transfers the management capabilities to the educational establishment so that it can gradually assume the running of the project without depending on external services.

To ensure the appropriation of the Eduinnova formative assessment proposal in the schools, the methodology defines a form of collaborative work that involves all of the different actors: school administrators (principals, etc.), technical support staff, teachers, and students. The administrators must provide the necessary technical and organizational framework and supervise the pedagogical quality of the planned activities; the technical support staff are responsible for maintaining the functionality of the mobile devices and associated software; the teachers are in charge of planning the activities, integrating them into the learning process, and guiding their execution in the classroom as mediators; and the students take on the role of activity protagonists, carrying out the actual collaborative work in groups of three.

The application of Eduinnova begins with the planning of a class, including a learning activity. The teacher prepares the activity with a specific objective in mind and then selects the desired content from a Web system accessed through the Internet. Since ICTs do not necessarily improve the quality of classroom instruction unless they are well integrated into the teaching curriculum (Keengwe, Onchwari, & Wachira, 2008), the methodology is implemented in conjunction with a 13,000-item database that covers mathematics and language subject matter from 3rd to 12th grade, science from 5th to 8th grade, and biology, chemistry, and physics from 9th to 12th grade.

Eduinnova's Web-based content management system allows the instructor to do the following:

- Review, edit, and modify the different questions organized by learning (i.e., subject) areas.
- Create questions and store them in the learning area they were defined for.
- Select the desired questions to create activities that can be stored for later use.

Once an activity is created, it is downloaded from the Web site to the teacher's machine.

To carry out an activity in a class, the teacher starts by handing out a mobile device—either a Pocket PC or an XP machine such as the Classmate

PC—to each student. To boot them up, each students must enter his or her identification number. A wireless network randomly distributes the class into groups of three (Nussbaum et al., 2009) and the students then physically separate into the groups so defined. At this point the teacher sends the activity to the students via the wireless network, and the actual collaborative work begins.

When the activity session ends, the monitoring information used by the teacher to mediate the activity groups (Figure 11.2) is stored by the teacher on his or her machine. It is automatically sent on to the Web site so that the teacher, the school principal, and other actors in the educational process can make use of it in decision making related to the progress of the course.

ANALYSIS OF THE EDUINNOVA PROJECT

Design of the Analysis

A study was carried out to ascertain the effects of the Eduinnova formative assessment methodology at the schools where the project was implemented. The first part of the study incorporated a mixed qualitative and quantitative design that aimed to capture the perceptions of the various actors involved in the project as they related to the implantation of the project and its ultimate impact on themselves, the students, and the school. The second part was intended to determine the final outcomes in terms of student learning. In addition to the actors' perceptions of the outcomes, this evaluation involved an analysis of students' scores on a standardized performance test known as SIMCE (Education Quality Measurement System), which is applied annually by the Chilean Ministry of Education to all children in the fourth grade.

The data for the qualitative aspects of the actors' perceptions were collected through interviews and focus groups and analyzed using the techniques of content analysis. In this method the importance of an idea in a text is indicated by the frequency of its appearance (Morse, 1995).

The content analysis consisted of the following steps:

1. The qualitative data were examined to find patterns in the actors' observations (open coding)
2. The patterns were grouped into general categories, thus yielding a hierarchical content scheme.

3. The significant categories were identified quantitatively by determining the themes that appeared most frequently.
4. Finally, the content of these categories was examined to identify perceptions and meanings.

The four categories defined by this analysis are context and general considerations, the teacher during the collaborative activity, the students during the collaborative activity, and classroom interactions. Further, on the Eduinnova project results and their analysis, these categories are each exemplified with quotations from the actor interviews and focus groups.

The quantitative data on the actors' perceptions were gathered via a structured survey and analyzed on the basis of responses to questions relating to four social variables: mutual support, listening to one other, giving others explanations, and resolving problems.

As for the evaluation of student learning outcomes, this consists of two analyses of students' scores on the above-mentioned national SIMCE test for each of three broad learning (subject) areas: language, mathematics, and natural and social sciences. The first analysis relates to the differences in scores between students at the schools participating in the project and students at other schools; the second one identifies differences where the participating students used Eduinnova in some but not all of the subject areas. The statistical technique used for these analyses was the *t* test for differences between means.

Definition and Selection of Sample

The study covered 23 municipal (state-subsidized) schools in the Chilean cities of Santiago and Antofagasta where the Eduinnova project was implemented. The number of actors involved in the project at these educational institutions are indicated by category in Table 11.1.

RESULTS OF THE ANALYSIS

Context and General Considerations

A range of different views were expressed by the teachers in the focus groups and interviews on what they understood by collaborative, cooperative, and group work. There was a consensus, however, on the

Table 11.1

ACTORS PARTICIPATING IN THE EDUINNOVA PROJECT

ACTORS IN PROJECT	TOTAL
Students	2,731
Teachers	272
Pedagogical coordinators	9
School principals	9
Technical support staff	9

complexities they faced in generating collaborative learning experiences in the classroom. They perceived that group and cooperative work were easier to achieve because in these modes the students work on their own before getting together to come up with a final product. This could be done even with homework assignments. Collaborative work, on the other hand, is understood, defined, and differentiated on the basis of interaction among group members, who need each other in order to attain a common objective. It is this positive interdependence and the generation of consensus and communication among students that is felt to be very difficult to achieve and generally does not occur. For example, one teacher from Providencia in Santiago reported that "I have always found it extremely difficult to do collaborative work with my students. I could say that collaborative work is sitting in groups, but that in itself is not in fact collaborative work."

When the students were asked about their preferences as between traditional classes and those using the collaborative dynamic, 80% said they preferred classes with portable technology and Eduinnova methodology.

The Teacher During the Collaborative Activity

The perception of the various actors was that with the Eduinnova methodology, a change occurs in the teacher's classroom role. Most of the teachers stated that whereas in their usual activities the frontal, expository style of instruction is predominant, with the implementation of the Eduinnova project they began adopting different styles and dynamics, combining the frontal or expository modes with others that emphasized a more participatory approach and generated a more symmetrical relationship with their pupils. The educator thus assumes the role of facilitator in which, instead of merely delivering knowledge, he or she facilitates the construction of knowledge among and from the students.

> I tend to give very frontal classes. . . . I explain, ask questions, taking the protagonist role. However, with Eduinnova, they [the students] are the ones who do everything, so in addition to monitoring I join a group and offer opinions. (Teacher, Peñalolén, Santiago)

In their roles as facilitators and monitors, the teachers stated that their relationship with the students was more symmetrical, allowing them to provide greater support and be more tolerant. They also took on the role of mediator, handling disagreements and conflicts among the students. This experience also validates the importance of their participation in the activity.

All of the actors involved in the implementation of the project (teachers, school principals, pedagogical coordinators, and technical support staff) referred to the methodology and the in-class management graphic tool (Figure 11.2) as elements that facilitated and assisted in the teaching task. They felt that the tool enabled them to follow the students' learning process during the collaborative activities and also determined the performance of the teachers, thus accounting for their positive views of the technology.

> Using the tool I observe the progress of the activity. If I see that a group is stagnating while another one is advancing, apparently at random, I go to the second one, sit down beside them, and if I find they're working well I go to the other group. If it turns out from the mistakes registering on the tool that they're just guessing at the answers, I suggest we go over together the work they're doing. (Teacher, Peñalolén, Santiago)

> Both the school administrators and pedagogical coordinators report the project's real potential for following the progress of teaching processes and supporting the task of teaching and creating new and better coordination machinery within the schools. (Focus Group report, Las Condes, Santiago)

As for the quantitative survey data, when the teachers were asked how they evaluated their own performance during the Eduinnova collaborative dynamic as regards the attention given to the students during the learning process, 96% said they devoted as much or more time to them than in traditional classes. The majority of teachers stated that they put in equal or greater effort compared to other classroom dynamics, and almost 80% commented that they played a mainly monitoring role during the activity, supervising groups and guiding them through the learning process. The majority of the students (63.3%) declared that they had received enough support from the teachers, while 36.6% indicated they would have liked to receive more.

The Students During the Collaborative Activity

All of the actors stated that with the collaborative dynamic, the students assumed an active role in the learning process given that the tool challenged them by requiring they record what they had learned, using their personal resources and skills to progress through the activities. Technical support staff in Las Condes, Santiago reported that "The students become the protagonists, they themselves feel that this is the case and are motivated and enthusiastic."

The teachers noted the appearance of leaders among the students, who made sure their groups advanced through the activity and did not fall behind. They took it upon themselves to manage their fellow group members and to involve them in the process. Another role reported by the teachers was that of the "student monitor," who supported not only his or her own group but the entire class as well. These students would sometimes finish before the other groups and would then go around the classroom giving assistance wherever it was needed. Other students emerged in roles taking charge of the computers, their maintenance and distribution, and so on.

The various actors also expressed the view that the methodology led to the development of various cognitive skills and clearly identified the existence of interaction among the students and its value.

> In the groups the students observe how their peers' thinking develops. If a student says that the answer is such-and-such, he or she will then have to say why. Giving arguments for one's responses is super important for the development of the student's thinking, and I saw this process occurring in various groups. (Pedagogical coordinator, Las Condes, Santiago)

The teachers further stated that the students were forced to interact and communicate, thus developing communication skills as they solved the problems they were given, even if they did not otherwise have good relationships.

> Some children may not get along well . . . or converse very much, but they are obliged to interact and communicate. So we don't have a situation in which friends keep to themselves within the groups and ignore the other members. They all interact with each other and the course itself functions much better. (Teacher, Peñalolén, Santiago)

Finally, the teachers found that the methodology promoted development of socioaffective skills required for collaborative work, such as respect and tolerance of one's classmates.

As for the quantitative analysis, Table 11.2 summarizes the teachers' responses to the survey on perceptions of the development of the four social variables (mutual support, listening to one other, giving others explanations, and resolving problems) in their students. The figures indicate the percentage of teachers who said that the variables showed more, the same, or less development under Eduinnova compared to the traditional lessons.

The students' responses to the same survey generally coincided with those of the teachers. They were less likely to say that the social variables showed greater development under the methodology, but the majority of them stated that all of the variables—mutual support, listening to one another, giving others explanations, and resolving problems—occurred to an equal or greater extent than with traditional methods. The students thus testified to a higher degree of interaction among themselves in a collaborative work environment mediated by portable technology.

Table 11.2

TEACHERS' PERCEPTIONS OF THE DEVELOPMENT OF SOCIAL VARIABLES AMONG STUDENTS DURING COLLABORATIVE ACTIVITIES AS COMPARED TO A TRADITIONAL CLASS

	MUTUAL SUPPORT	LISTENING TO ONE ANOTHER	GIVING OTHERS EXPLANATIONS	RESOLVING PROBLEMS
More	82%	65%	79%	66%
Equal	17%	33%	20%	32%
Less	1%	3%	1%	2%

Classroom Interactions

Since the students are divided into groups randomly, the members of a given group are sometimes not accustomed to interacting with each other. As they get to know each other, they often discover new aspects of their classmates' personalities.

> When one of the children told me it was the first time he'd been together with two of the others since fifth grade, I asked him how that could be since he was now in eighth grade, and he said, "they just never talked to me before but now with the Pocket PCs we talked" and I saw that the technology isn't so cold after all. (Teacher, Providencia, Santiago)

The project actors perceived that during the activity the students helped each other, gave explanations to others, and displayed much collaborative behavior that would normally be difficult to achieve.

> One learns to spot the students who have problems working in a group due to their character or personality, or because they don't like the other members. [With this system] they are forced to participate. There were groups where the students wouldn't even look at each other and starting squabbling, but they soon realized they were falling behind the other groups and had no choice but to mix, work together, participate and come to agreements. (Teacher, Las Condes, Santiago)

And in the words of a Las Condes pedagogical coordinator, "It's an extremely powerful tool for communication between them and the teacher."

All of the participating actors agreed that the activity motivated the students to perform the tasks and learn, making them more active and concentrated on what they were doing, and that they found the methodology attractive in and of itself. Another coordinator reported that "They are motivated and [the activity] allows them to be competitive. . . . Suddenly they all want to improve and their self-esteem increases."

The teachers also reported that, with the methodology, they perceived a positive environment in the classroom. It is true that when it is first implemented there is a certain amount of disorder as the children divide up into groups, but once the groups are formed, the activity runs smoothly in a relaxed and positive atmosphere.

> The class holds the children's attention and keeps them involved. With the incorporation of the technology, discipline in the classroom has improved

> because the students are anxious to do their work and carry out the tasks together with their classmates, so they simply forget about being disorderly. (Teacher, Huechuraba, Santiago)

The teachers noted that under Eduinnova they can keep the whole class under control even when they are busy mediating a specific group, which would be impossible in a conventional lesson.

> It's not like a frontal class where all the children are looking at you, and you're always having to deal with some students who are talking, or make sure others are learning or not causing a disturbance or paying attention. Here, each child is busy, and if they have questions they raise their hands. You can deal with one group without the others taking advantage to misbehave. The activity is interesting for them so they're all involved in it. (Teacher, Peñalolén, Santiago)

As regards the quantitative perceptions survey, the majority (75%) of the teachers observed that the students shared more knowledge under the activity dynamic than in other learning contexts, showed more interest in doing a good job with the collaborative methodology, and valued positively the fact that the groups were defined randomly.

Student Learning Outcomes

As noted earlier, the impact of Eduinnova on actual learning outcomes was determined in terms of both the perceptions of teachers and students gathered by a structured survey and an analysis of the results of the national SIMCE (2007) test given to fourth grade students.

Perceptions of Teachers and Students

The perception of all of the actors in the project was that student learning attainment occurred under the collaborative methodology. The report of a technical supporter said that "The strengths of the [collaborative] work with PDAs in the classroom is that it improves learning. The students' concentration and relations between them are better, time is optimized, discipline is stronger and the burden on the teachers is lighter."

The majority of the teachers and students also expressed the view that Eduinnova showed better results than traditional classroom

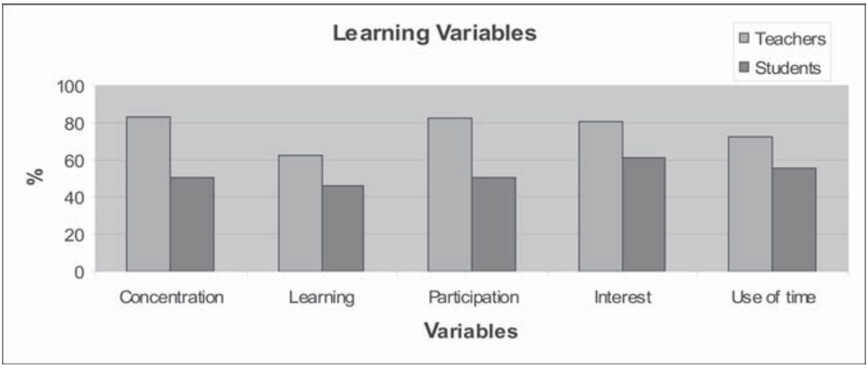


Figure 11.3 Actors participating in Eduinnova project.

methods on five variables considered to be related to significant learning. This is illustrated in Figure 11.3, which indicates the percentages of teachers and students who reported this perception for concentration, learning, participation, interest, and use of time.

In particular, the survey results indicated that almost half of the students (45.9%) felt they learned more with the project methodology, while a similar proportion (45.9%) said they learned the same amount as in traditional classes and 8.2% reported that they learned less. The teachers' responses, on the other hand, revealed that 62% believed the students learned more with the project methodology, 35% claimed they learned the same amount, and only 2% reported that they learned less.

Analysis of SIMCE Test Scores

Whereas the foregoing analyses[1] of perceptions included the results for all 23 schools participating in the Eduinnova project, the evaluation of the learning test results is based on only 5 of them, all located in the Antofagasta urban area in northern Chile. The other 18 schools were excluded from this part of the study for a number of reasons. First, 12 of them had not yet joined the project when the SIMCE (2007) test was given. Second, 5 of the 6 remaining ones constituted the entire set of municipal (state-subsidized) institutions in a single district of Santiago and were therefore not comparable to the Antofagasta schools. Finally, at the 6 Santiago establishments the number of grades in which no courses used Eduinnova was relatively low (105 students in three grades, two at the same school), making it impossible to derive estimates for comparisons between students with a sufficient degree of validity. This problem will be eliminated in future studies, in which the participating subjects will be defined more strictly to ensure the required level of compatibility for making comparisons.

Of the total of 12 courses at the five schools, 11 took part in the project; all of the 11 courses participated in mathematics, 6 in language, and 9 in natural and social sciences. The number of students thus involved was 487. In the analyses that follow, their SIMCE scores are compared with those of the control group, which consisted of the approximately 2,670 students attending the various subsidized schools in Antofagasta where the Eduinnova project was not applied. For both groups, the scores analyzed were the averages on each of the three subject areas.

Comparison of SIMCE Test Scores by Subject Area With and Without Eduinnova

The first SIMCE score analysis compared the students from the participating and non-participating (control group) schools on their results in the three SIMCE subject areas. A *t*-test of the scores found statistically significant differences between the two groups in all three subjects, implying that in each case the group using Eduinnova performed better. The details are given here in Table 11.3.

Comparison of SIMCE Test Scores With Eduinnova Used in Some Subject Areas

The second SIMCE score analysis consisted of two sets of comparisons aimed at identifying performance differences in cases where students used Eduinnova in some but not all subject areas. It was expected that such students would have higher scores than the control group mean in subjects where the methodology was applied and similar scores where it was not.

Since the test scores for the three subject areas are not directly comparable with each other, the results discussed in what follows for each subject were centered around the mean for all Antofagasta schools and weighted by the corresponding standard deviation.

The first set of comparisons contrasted students who used the methodology in mathematics but not language classes with the control group (Table 11.4). The Eduinnova group diverged from the Antofagasta mean on the mathematics test by 0.34 standard deviations, a difference that is statistically significant ($p < 0.05$; $1 - \alpha = 95\%$). They did not differ

Table 11.3

COMPARISON OF SIMCE SCORES BY USE OF EDUINNOVA AND SUBJECT AREA

SUBJECT AREA	STUDENT GROUP	NO. OF STUDENTS	MEAN	DIFFERENCE	SIGNIFICANCE	COHEN'S D
Mathematics	Not using Eduinnova	2,593	219.86	17.17 (7.8%)	$p < .05$	0.32
	Using Eduinnova	433	237.031			
Language	Not using Eduinnova	2,756	231.97	6.72 (2.8%)	$p < .05$	0.13
	Using Eduinnova	233	238.69			
Natural and social sciences	Not using Eduinnova	2,669	225.43	11.03 (4.8%)	$p < .05$	0.23
	Using Eduinnova	354	236.46			

Note: t test with $1-\alpha = 95\%$.

from the nonparticipating Antofagasta mean on the language test (.05 standard deviations, not statistically significant: $p > 0.05$; $1 - \alpha = 95\%$). This implies that the students performed better only in the subject area where Eduinnova was applied.

The second set of comparisons contrasted students who used the methodology in natural and social science but not language classes with the control group (Table 11.5). The group that worked under Eduinnova scored .34 standard deviations higher than the Antofagasta mean for natural and social sciences and .22 standard deviations higher than the nonparticipating Antofagasta mean in language. This demonstrates that the students who had any experience with the collaborative methodology were above the Antofagasta mean in both subject areas and clearly performed better in the subject where the methodology was applied.

As regards the language test, no comparisons were possible because each grade which used Eduinnova in language classes also used it in mathematics and in natural and social sciences.

DISCUSSION

The presented face-to-face technology-supported small group collaborative learning strategy follows the constructivist ideas of Activity Theory (Engeström, 1987), where students are participants of their learning, identifying three pillars:

1. Collaboration: Members of a group work together on the same objective. Each group member is responsible for his/her own work, role and learning effort, and each one constructs, prior to the collaborative discussion, his/her own vision of the problem. Thus, all group members are a source of information for building the common answer as they help each other reach the common goal.
2. Reflection: The experience of each student is reflected in his or her exploration of the reasoning and viewpoints of every other one, and all of them thus become experts in their own learning. In this way a shared understanding of the task is constructed, enabling the students to exchange opinions, negotiate, and construct an answer together. A space is therefore required in which they can propose and defend the solutions they arrive at and ask their peers to clarify and justify their views (Kruger, 1993). Students exchange opinions through face-to-face social interactions and

Table 11.4

COMPARISON OF TEST SCORES: EDUINNOVA USED IN MATHEMATICS ONLY

MATHEMATICS SCORE	NO. OF STUDENTS	MEAN	*SD*	DIFFERENCE	*SD* DIFFERENCE	SIGNIFICANCE
Without Eduinnova	2,593	219.86	52.58	18.43 (8.3%)	0.35	$p < 0.05$
With Eduinnova in math but not language	202	238.3	54.19			
LANGUAGE SCORE	**NO. OF STUDENTS**	**MEAN**	***SD***	**DIFFERENCE**	***SD* DIFFERENCE**	**SIGNIFICANCE**
Without Eduinnova	2,547	231.75	49.94	2.99 (1.2%)	0.05	$p > 0.05$
With Eduinnova in math but not language	209	234.74	51.43			

Note: t test with $1-\alpha = 95\%$.

Table 11.5

COMPARISON OF TEST SCORES: EDUINNOVA USED IN NATURAL AND SOCIAL SCIENCES ONLY

NATURAL AND SOCIAL SCIENCES SCORE	NO. OF STUDENTS	MEAN	*SD*	DIFFERENCE	*SD* DIFFERENCE	SIGNIFICANCE
Without Eduinnova	2,669	225.43	45.6	15.94 (7.07%)	0.34	$p < 0.05$
With Eduinnova in natural and social sciences but not language	121	241.37	48.58			
LANGUAGE SCORE	**NO. OF STUDENTS**	**MEAN**	***SD***	**DIFFERENCE**	***SD* DIFFERENCE**	**SIGNIFICANCE**
Without Eduinnova	2,631	231.45	49.95	11.48 (4.9%)	0.22	$p < 0.05$
With Eduinnova in natural and social sciences but not language	125	242.94	51.17			

Note: t test with $1-\alpha = 95\%$.

make compromises in order to build a consensus answer (Zurita & Nussbaum, 2007). Small groups whose members are in close physical proximity facilitate the social interactions and consensus-building required to construct an answer together (Adams & Hamm, 1996; Johnson & Johnson, 1999).

3. Significance: Learning has to be with a meaning. The tasks chosen for peer collaboration must be structured so that the students are obliged to work together cooperatively toward a common goal. Through joint actions and verbal explanations, the denotation of the task is discovered.

The perceptions of the actors in the Eduinnova project as presented and analyzed previously allow us to make a number of important observations in terms of the four categories defined by the content analysis as set out below.

Context and General Considerations

The actors distinguished between collaborative learning and group work, noting that the first of these is more difficult to achieve in the classroom and demands interaction between students to attain a common objective. They found that the Eduinnova project facilitated the achievement of collaborative learning. They also felt proud and privileged to be at the schools that participated in the project and valued both the project experience and the support provided by the implementation team.

The Teachers' Experience in the Collaborative Process

The teachers perceived that with the application of the Eduinnova methodology, their role in the classroom changed from being expositors of learning to being facilitators of it. They also played the parts of mediator between the students and monitor of the teaching–learning process, supervising the student groups and accompanying them as they learned.

All of the actors perceived that teaching performance was facilitated by the methodology, which provided the tools for supporting and assisting in the pedagogical task. A characteristic that was particularly stressed was the ability to follow the students' learning process and receive immediate delivery of activity results. This made it possible to intervene where necessary and provide feedback to students, thus improving learning effectiveness.

The Students' Experience in the Collaborative Process

The actors perceived that the students also experienced a change of roles during the collaborative activities, becoming active agents in their own learning processes. Leaders and monitors emerged spontaneously. The actors also reported that the students developed both cognitive and socioaffective skills.

Classroom Interactions

All of the actors mentioned that collaborative learning took place during the dynamic supported by the Eduinnova methodology. They perceived it as a tool that facilitates this collaborative dynamic and promotes tolerant and committed interaction between group members. Teachers and school principals valued positively the fact that the groups were formed in random fashion, as this promoted both integration between students who were previously not acquainted and learning between peers. This in turn encouraged interaction among them.

In addition, the actors all spoke of the motivation that was generated among the students by Eduinnova, thanks to its novelty, as well as the impacts it had on them, especially those related to interaction with their classmates.

Finally, the actors also declared that there was a genuine change in the classroom learning dynamics, creating a different climate and a new type of discipline that required conversation, discussion, and the student to act as a protagonist. All of the foregoing resulted in a better integration of the students, which the groups then carried with them to their other courses.

The results and analyses presented above also suggest how the Eduinnova methodology may comply with the principles for effective teaching and learning defined by the Teaching and Learning Research Program of the UK-based Economic and Social Research Council in a report entitled *Improving Teaching and Learning in Schools* (ESRC and TLRP, 2006). Five of the ten principles outlined there are listed below, with a brief statement on how they are supported by face-to-face technology-supported small-group collaborative learning:

- Equip learners for life. Social abilities facilitated by small-group work empower students as members of a working team and society. Face-to-face small-group collaborative learning inculcates social and communicative abilities while also teaching the curricula, as shown by the analyses presented above.

- Scaffold learning. Teachers need tools to diagnose learning difficulties, a task that can be performed by the in-class management tool presented here (Figure 11.2). Also, activities should provide intellectual, social, and emotional support to help learners move forward in their learning so that once these supports are removed, the learning is secure. Small-group formative assessment promotes intellectual abilities while simultaneously developing social and communicative abilities facilitated by the face-to-face small-group interaction and technology support.
- Assessment must be congruent with learning. The use of assessment to adapt teaching to meet learning needs is defined as formative assessment. *Assessment should be designed and implemented with the goal of achieving maximum validity both in terms of learning outcomes and learning processes. It should help to advance learning as well as determine whether learning has occurred* (TLRP & ESRC, 2006). The in-class management tool provides this.
- Promote the active engagement of learners. The promotion of learners' independence and autonomy through the acquisition of a set of learning strategies and practices that enables them to be agents in their own learning is an objective that was achieved in this project. Each student must take responsibility for his or her work through self-assessment when answering a question and then through peer assessment as the group searches for agreement on a single answer.
- Foster individual and social processes and outcomes. As we have seen, the strategy presented in this study encourages the building of relationships and communication with others through the learning objectives, which assist in the mutual construction of knowledge in order to enhance the achievement of the group and its peer members.

CONCLUSION

Formative assessment plays a role in knowledge practice and reinforcement through the evaluative practice of posing questions. Students should develop the capacity to monitor the quality of their own work, understand what high quality work is and where their own work stands in relation to it, and develop criteria for modifying their own work (Sadler, 1989). With the technology-supported face-to-face small-group formative assessment strategy described in this paper, the students' work is monitored through peer evaluation and then by the in-class management graphic tool's group evaluation. The teacher's role, supported by the graphic tool, is to mediate when and where such action is required. Finally, it was demonstrated by actual classroom application that implementation of the Eduinnova strategy leads students to modify their group work habits, thus bringing about an improvement in their social abilities and in learning.

ACKNOWLEDGMENT

This chapter was partially funded by the project Centro de Estudios de Políticas y Prácticas en Educación, CIE01—CONICYT.

NOTE

1. The present research used the sources of databases from SIMCE of the Chilean Ministry of Education. The corresponding authors appreciate the access to the information given by the Ministry of Education. All results presented in this study are the responsibility of the authors and in no way to be ascribed to this institution.

REFERENCES

Adams, D., & Hamm, M. (1996). *Cooperative learning: Critical thinking and collaboration across the curriculum* (2nd ed.). Springfield, IL: Charles C Thomas.

Black, P. (2005). Formative assessment: Views through different lenses. *Curriculum Journal, 16*(2), 133–135.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 16*(1), 9–21.

Black, P., McCormick, R., James M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education, 21*(2), 119–132.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–73.

Conlon, T., & Simpson, M. (2003). Silicon Valley versus Silicon Glen: The impact of computers upon teaching and learning: A comparative study. *British Journal of Educational Technology, 34,* 137–150.

Cortez, C., Nussbaum, M., Rodríguez, P., López, X., & Rosas, R. (2005). Teacher training with face to face computer supported collaborative learning. *Journal of Computer Assisted Learning, 21,* 171–180.

Deaney, R., Ruthven, K., & Hennessy, S. (2006). Teachers' developing "practical theories" of the contribution of information and communication technologies to subject teaching and learning: An analysis of cases from English secondary schools. *British Educational Research Journal, 32*(3), 459–480.

Dillenbourg, P. (Ed.). (1999). *Collaborative learning: Cognitive and computational approaches.* Oxford, UK: Pergamon.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research.* Helsinki: Orienta-Konsultit.

Enlaces. (2008). Plan Tec—Enlaces al Bicentenario. Santiago, Chile: Centro de Educación y Tecnología del Ministerio de Educación. Retrieved August 1, 2008, from http://portal.enlaces.cl/?t=44&i=2&cc=808&tm=2

ESRC (Economic and Social Research Council) & TLRP (Teaching and Learning Research Program). (2006). Teaching and learning in schools. Retrieved April 29, 2008, from http://www.tlrp.org/pub/documents/TLRP_Schools_Commentary_FINAL.pdf

Hayes, D. N. A. (2007). ICT and learning: Lessons from Australian classrooms. *Computers & Education, 49*(2), 385–395.

Inkpen, K. (1999). Designing handheld technologies for kids. *Personal Technologies Journal, 3*(1&2), 81–89.

Johnson, D. W., & Johnson, R. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book Co.

Johnson, D. W., & Johnson, R. T. (1999). *Learning together and alone. Cooperative, competitive, and individualistic learning.* Boston: Allyn & Bacon.

Keengwe, J., Onchwari, G., & Wachira, P. (2008). The use of computer tools to support meaningful learning. *AACE Journal, 16*(1), 77–92.

Kruger, A. C. (1993). Peer collaboration: Conflict, cooperation or both? *Social Development, 2,* 165–182.

Kumar, V. (1996). *Computer-supported collaborative learning: Issues for research.* Paper presented at the 8th Annual Graduate Symposium on Computer Science, University of Saskatchewan.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary student's computer-supported collaborative learning. *Learning and Instruction, 13,* 487–509.

Nussbaum, M., Alvarez, C., McFarlane, A., Gomez, F., Claro, S., & Radovic, D. (2009). Technology as small group face-to-face collaborative scaffolding. *Computers and Education, 52*(1), 147–153.

Nussbaum, M., Furman, A., Feuerhacke, A., Radovic, D., Gomez, F., & Lopez, X. (2007). Comparative study of peer learning mediated by interconnected PCs and PDAs. mLearn Meeting Proceedings Melbourne, Australia, October 16–19, pp. 186–191.

Morse, J. (2003). *Critical issues in qualitative research methods*. Universidad de Antioquia, Medellín, Colombia.

Postholm, M. B. (2007). The advantages and disadvantages of using ICT as a mediating artefact in classrooms compared to alternative tools. *Teachers & Teaching, 6*(13), 587–599.

Reynolds, D., Treharne, D., & Tripp, H. (2003). ICT: The hopes and the reality. *British Journal of Educational Technology, 34*(2), 151–167.

Robertson, J. (2003). Stepping out of the box: Rethinking the failure of ICT to transform schools. *Journal of Educational Change, 4,* 323–344.

Roschelle, J., Pea, R., Hoadley, C., Gordon, D., & Means, B. (2001). Changing how and what children learn in school with computer-based technologies. *The Future of Children, 10*(2), 76–101.

Roschelle, J., Rosas R., & Nussbaum, M. (2005, July). Towards a design framework for mobile computer-supported collaborative learning. *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning*, Taiwan, pp. 520–524.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144.

Salomon, G., & Almog, T. (1998). Educational psychology and technology: A matter of reciprocal relations. *Teachers' College Record, 100*(1), 1–20.

Savill-Smith, C., & Kent, P. (2003). *The use of palmtop computers for learning. A review of the literature.* London: LSDA. Retrieved from www.LSDA.org.uk

SIMCE. (2007). *Data Base*. Santiago, Chile: SIMCE, Ministerio de Educación.

Soloway, J. E., Norris, M. C., Jansen, R. J., Krajcic, R. M., Fishman, B., & Blumenfeld, P. (2001). Making palm-sized computers the PC of choice for K-12. *Learning and Leading With Technology*, *28*(7), 32–34, 56–57.

Tait, B. (1997). Constructive Internet based learning. *Active Learning, 7,* 3–8.

Vygotsky, L. (1930/1978). *Mind in society.* Cambridge, MA: Harvard University Press.

Watson, D. (2001). Pedagogy before technology: Re-thinking the relationship between ICT and teaching. *Education and Information Technologies, 6*(4), 251–266.

Wood, D. O., & Malley, C. (1996). Collaborative learning between peers: An overview. *Educational Psychology Practice, 11*(4), 4–9.

Zurita, G., & Nussbaum, M. (2004a). A constructivist mobile learning environment supported by a wireless handheld network. *Journal of Computer Assisted Learning, 20,* 235–243.

Zurita, G., & Nussbaum, M. (2004b). Mobile computer supported collaborative learning. *Computers & Education, 42*(3), 289–314.

Zurita, G., & Nussbaum, M. (2007). A conceptual framework based on activity theory for mobile CSCL. *British Journal of Educational Technology, 38*(2), 211–235.

This page intentionally left blank

12

Folk Pedagogy and Cultural Markers in Teaching: Three Illustrations From Chile

DAVID D. PREISS

The purpose of this chapter is to show the main results of an ongoing line of research focused on understanding teaching as a cultural practice. In so doing, I attempt to tap into teachers' folk pedagogies—that is, the implicit models of pedagogy shared by theorists, educators, and children (Bruner, 1996; Olson, 2003; Olson & Bruner, 1996; Torff & Sternberg, 2001). Specifically, the focus of this work is on the instructional principles that, in Chile, give cultural and personal meaning to the task of teaching language and mathematics and how these principles translate into real teaching practices.

My interest in folk pedagogy was triggered by two facts: first, the growing interest in psychology for schooling and the cognitive consequences of schooling; second, the polemical nature of the contemporary interpretation of educational quality issues as caused by a crisis in teaching.

A research grant of the Pontificia Universidad Católica de Chile's VRAID and a dissertation fellowship of the Yale Graduate School supported the elementary school study. Grant Number 11060389 from FONDECYT supported the two middle-school studies. Preparation of this paper was supported by the Center for Research on Educational Policy and Practice, Grant CIE01-CONICYT. Special thanks are given to the Centro de Medición Mide UC, which granted the author access to all the databases and material used for this study. I am grateful to my research assistants Susana Valenzuela and Ana María Espinoza, whose diligent work helped to advance the middle-school studies presented here.

Today, there is a large amount of empirical evidence indicating that schooling affects cognitive development, particularly those dimensions of cognitive development measured by conventional tests of intelligence (Blair, Gamson, Thorne, & Baker, 2005; Ceci, 1991; Cole, 2005; Olson, 1994, 2003, 2005). In addition to the cognitive gains associated with schooling, scholars have also observed its impact on well-being, health, and economic development (Healy Côté, Helliwell, Field, Centre for Educational Research and Innovation, & Organisation for Economic Co-operation and Development, 2001). Thus, reports on school achievement receive increasing attention from the general public everywhere, especially when these reports deliver news not entirely positive. Prompted in many cases by international comparative studies, educational crises of some sort are reported in countries differing in the nature of their political arrangements, their economic resources, and even the levels of achievement they attain (Gorard, 2001; Haye & Pacheco, 1995; Stigler & Hiebert, 1999; Takayama, 2008; Takayama & Apple, 2008).

This widely shared concern for educational quality has recently been reframed as a concern for teaching quality. Different lines of research have contributed, deliberately or by coincidence, to turn public attention to teachers. Particularly influential were the TIMSS 1995 and 1999 video studies, which focused on cultural differences in instruction and their possible links to international differences in student achievement (Givvin, Hiebert, Jacobs, Hollingsworth, & Gallimore, 2005; Stigler & Hiebert, 1999). Complementing and in some cases inspired by the TIMSS studies, a new emphasis on comparative pedagogical studies began to emerge in different countries, either by the publication of large-scale cross-cultural studies (Alexander, 2000, 2001) or single-culture studies (e.g., Loera, 2006; Preiss, 2005, 2009). Additionally, public policy sources began to address the issue of teaching as playing a key role in educational achievement. For instance, one public policy report, which has influenced public discussion on education in nations as diverse as Chile and the United Kingdom, regards teachers as the main drivers of variations in student learning in the classroom (Barber & Mourshed, 2007). Consequently, issues of educational accountability have increasingly become issues of teaching accountability, and Chile is not an exemption to this trend (Avalos, 2004).

Teaching quality can be approached from multiple perspectives. Pursuing a definition of good-quality teaching, Fenstermacher and Richardson (2005) made a distinction between *good* teaching and *successful* teaching that is worth considering here: "Good teaching is grounded in the task sense of teaching, while successful teaching is grounded in the achievement sense of the term" (Fenstermacher & Richardson, 2005, p. 189). Good teaching is teaching that fulfills high standards for both subject matter content and delivery. More specifically, good teaching is in harmony with instructional principles that are morally and rationally defensible. Successful teaching is primarily concerned with delivering the intended learning; for instance, to provide students with the skills they need to achieve high scores on standardized tests.

This chapter is concerned with good teaching. It is my belief that, to understand potential strengths and deficits in teaching performance, we must go beyond performance-assessment measures, either at the student or the teacher level, and look at teacher practices from the privileged vantage point of teacher folk pedagogies. Thus, my focus is not on the short-term strategies teachers can use to raise scores in standardized tests but on the instructional principles that—particularly in Chile—give cultural and personal meaning to the task of teaching.

The plan of the chapter is as follows. Because the theories here applied have been primarily advanced in North America (the United States and Canada), I first comment on the sociocultural issues related to the testing of current North American instructional theories in Chile. In so doing, I assert that one must proceed with caution in interpreting data using theories foreign to the context where those data were collected—particularly when this interpretation is associated with performance assessment. Second, I explore commonalities between Chile and the context of origin of the educational theories here employed. I assert that these commonalities make the application of these theories to the Chilean context not only feasible but also relevant. Third, I summarize contemporary definitions of folk pedagogy and advance a tentative operationalization, which I contrast with international evidence. Finally, using empirical data, I illustrate the folk pedagogies that prevail in Chile. These illustrations are based on classroom observations in Chile of elementary and middle-school teaching of language and middle-school teaching of mathematics.

SOCIOCULTURAL ISSUES AT STAKE

By assessing the pertinence of the theories of folk pedagogy to Chilean education, my work tests, in Chile, the applicability to and validity of current North American instructional theories. It is rare for these theories to be tested outside the context in which they were created. Yet they have to be tested in cultural milieus other than their own if we are to investigate what aspect of these theories, if any, can be considered part of a universal theory of teaching.

Values and beliefs that originated in highly literate, mostly individualistic, and mostly Protestant communities permeate core aspects of dominant educational theories. These theories rest on an individualistic pedagogy, which views learning as the development of higher-order cognitive skills and expert knowledge by the independent learner (see, for instance, Bransford, Brown, & Cocking, 1999). Thus, when context is considered, most of these theories put context in relation to the needs of the individual learner. In addition, two concepts central to contemporary research on education—metacognition and self-regulation (Dinsmore, Alexander, & Loughlin, 2008; Olson & Astington, 1993)—are highly attuned to the cultural backgrounds of these communities. On the one hand, there is a good fit between the weight given to "thinking about thinking" in contemporary educational theories and the metacognitive practices characteristic of literate cultures (Olson, 1994). On the other hand, the notion of self-regulation is highly compatible with the emphasis on self-control that is characteristic of Protestant cultures (Weber, 2001).

In contrast, Latin American cultures are not markedly literate or individualistic. In fact, Latin American thinking styles are closer to the literary than the scientific (Preiss & Strasser, 2006). A metacognitive style in speech and thinking is presumably less prevalent or less developed. Furthermore, Latin American religiosity has developed along different lines. Rather than privileging individual consciousness, Latin American Catholicism gives weight to public rituals (Morandé, 1984; Paz, 1985); therefore a metacognitive orientation is not characteristic of Latin American cultural practices. This basic orientation to the public expression of beliefs may find a classroom correlate in the observed adherence of language teachers in Chilean elementary schools to produce activities involving public play among students (Preiss, 2005). Therefore the absence of certain behaviors in the classroom and the presence of others may result not only from teaching deficits but also from more widely spread cultural patterns.

COMMONALITIES IN VIEWS OF LEARNING AND INSTRUCTION

Notwithstanding these cultural differences, there are relevant commonalities between Chilean views of learning and instruction and those that dominate in North America. These commonalities make the application of contemporary North American educational theories to the Chilean context not only useful but also relevant.

Researchers have documented two main attitudes to education in Chile: an *enlightened* approach and a *factory* approach (Haye & Pacheco, 1995). The former originated in the early decades of the 19th century, when Chile won its independence from Spain; the latter developed as the Chilean educational system grew in complexity and increased its recruitment levels. The main aspiration of the enlightened approach was to establish the cultural basis of the new society: the role of public education was to produce good citizens who adhered to the ideals of the Enlightenment, which inspired Chile's movement toward independence (Haye & Pacheco, 1995). The enlightened approach prevailed throughout the 19th century and was revitalized at the end of the century, when the German model of education was promoted by the state. As the educational system grew and state initiatives made elementary education more accessible for the masses, the influx of positivism began to change the state's approach to education. Thus the educational system moved toward adopting the standards of a "factory model of education." Throughout the 20th century, the practical needs of the developing Chilean educational system coexisted (and occasionally clashed) with the enlightened ideals that had previously driven the establishment of public education (Haye & Pacheco, 1995; Nuñez, 2002).

Education in North America was built on two historical sources; in some respects, it parallels the approaches we described for the Chilean case. The North American sources are the 19th-century classic model of education, which emphasized the rote learning of exemplary texts such as the Bible, and the 20th-century "factory model of education," which spawned mass education. As defined by Rogoff, Matusov, and White (1996), the factory model of education has as its main purpose the functional education of the masses. It imitates factory efficiency in its emphasis on standardized testing of the product. Both of these models shared a view of the teacher as an authority figure whose role it is to convey the right material to the children as prescribed either by the standards of religion or the standards of the state. Both of these models are based on a "transmission theory" of learning and are examples of what Rogoff and collaborators call *adult-run* instructional models: "The teacher's job in the adult-run model is to prepare the knowledge for transmission and to motivate children to make themselves receptive" (Rogoff et al., 1996, p. 393).

Although it goes beyond the goals of the present chapter to discuss the origin and nature of these commonalities, I would like to offer a hypothesis about the rapid adoption of the factory model of education by these two societies. In comparison with European or East Asian countries, the cultural foundations of both Chile and North America are relatively new. Consequently, as mass education was rooted in still fresh national cultures, the bureaucratic dimension of mass education could rapidly erode previously existing ways of educating. Unlike the situation in these relatively new countries, the establishment of a system of mass education in countries possessing older cultural traditions, such as Japan, may have involved an amalgamation of the universal practices of schooling with previously settled local practices of cultural transmission and educating. And that is why, during the nineties, teaching practices in Germany and especially Japan looked so distinctively "cultural" to the North American interpreters of the classroom videos taken in the context of the TIMSS initiative (Stigler & Hiebert, 1999).

VARIETIES OF PEDAGOGY

Adult-run pedagogies, such as those dominating Chilean and North American education, are commonly contrasted with student-run pedagogies (Rogoff et al., 1996). This contrast inherits a long-term differentiation between traditional and progressive pedagogies. Progressive pedagogies have also been called *constructivist* because of their emphasis on knowledge-construction processes at school (Mayer, 1998). The distinction between traditional and progressive teaching has driven much of contemporary research on instruction (Olson & Torrance, 1996). Although influential, the gross distinction between traditional and progressive—or any of the analogue distinctions—does not discern the diversity of pedagogical approaches teachers take in the classroom. The construct of folk pedagogy, however, is designed to be attentive to variations in practice.

Different folk pedagogies are based on different folk psychologies. Folk psychologies are the commonsense ways in which people understand subjective experiences (Bruner, 1990). Likewise, folk pedagogies are the commonsense ways through which educational agents understand the experiences of teaching and learning (Bruner, 1996; Olson, 2003; Olson & Bruner, 1996). In the same way that implicit theories of intelligence inform more formalized theories of intelligence (Sternberg, 1985a, 1985b, 1990), folk pedagogies inform more formalized theories of learning.

Thus teacher folk pedagogies can be added to the other three components of teacher knowledge identified by Shulman (1986): subject matter knowledge, pedagogical knowledge, and curricular knowledge. Folk pedagogies are evidently not subject matter or curricular knowledge, nor are they pedagogical knowledge, because they are not explicit pedagogies. Because they have all the attributes defining tacit knowledge—procedural structure, high usefulness, and low environmental support for acquisition—folk pedagogies can be seen as a key component of teachers' tacit knowledge (Sternberg et al., 2000). Indeed, folk pedagogies are more concerned with how to teach than with the question of what teaching is; they help teachers solve the practical problems they face within their classrooms, and they are for the most part shaped by the everyday practices of teachers and their previous practical experience as learners. In consequence, folk pedagogies are not merely an additive form of knowledge but constitute an organized representation of the instructional process.

Capitalizing on the distinction between traditional and progressive pedagogies, Olson and Bruner (1996) make a general distinction between externalist and internalist folk pedagogies, which, in turn, they subdivide into four categories. They do not provide a clear picture of how these folk pedagogies are implemented by teachers and do not propose a strict operational definition of the folk pedagogies they describe. In order to bridge theory and empirical research, I advance the following operationalization:

1. An internalist folk pedagogy focused on the *social construction of knowledge* involves extensive mentalistic and elaborative dialogical talk. Lessons are structured using a collaborative format, allowing students to play a central role while working on an instructional task. Teachers use a moderate number of follow-up communicative moves during the lesson.
2. An internalist folk pedagogy with a focus on the *cultural elaboration of knowledge* differs slightly. It involves extensive mentalistic and elaborative dialogical talk, amplified by a heavy use of teacher follow-up communicative moves. Lessons are structured using a negotiated format so that both teacher and students play shared roles while acquiring meaningful knowledge according to standards set by different cultural sources, including the teacher.

3. In contrast, an *externalist folk pedagogy*, which is focused on the transmission of content, involves more objective and informative talk and is regulated by a moderate use of evaluative teacher follow-ups. Lessons are canonically structured so that the teacher plays a key role while delivering content knowledge.
4. An *externalist folk pedagogy*, which is focused on skills training, involves pragmatic communicative exchanges regulated by an elevated use of evaluative teacher follow-ups. Lessons are canonically structured so that the teacher plays a key role while demonstrating procedural skills.

These operational definitions are schematized in Table 12.1.

THE EMPIRICAL STUDY OF FOLK PEDAGOGIES

Folk pedagogies can be approached empirically in two ways. One way is to study the folk pedagogies manifest in teachers' beliefs, using conventional survey methods or interviews. The second way is to infer folk pedagogies from the teaching practices they inform, using convergent and relatively indirect methods such as video surveys. These two alternatives roughly match the two dimensions that have been identified as the two major domains of research on teachers: teacher cognition and teachers' actions (Clark & Peterson, 1986). However, research on teaching actions has commonly been behavioristic. Its goal has been to relate certain teacher behaviors to student achievement. This study takes a different route, aiming to infer teachers' implicit pedagogies (teacher cognition) from observable behavior (teaching patterns).

To study instructional patterns from the point of view of folk pedagogies, we must look at those aspects of teaching practice that most influence the learning process. Among these features we must select those that allow a comparison with the knowledge accumulated in previous research. The following two dimensions of teaching fulfill these requirements.

Teacher Talk

Scholars working on teacher talk deem two dimensions to be critical for student learning: the pedagogical use of metacognitive content and initiate–respond–follow-up sequences (hereafter IRF sequences;

Table 12.1

OPERATIONAL DEFINITIONS OF FOLK PEDAGOGIES

FOLK PEDAGOGY	INTERNALIST		EXTERNALIST	
Emphasis	Social-constructionist	Culture- centered	Content transmission	Repeated practice
Talk	Dialogue	Dialogue	Monologue	Pragmatic exchange
Follow-up use	Moderated	High	Moderated	High
Lesson structure	Student-run	Negotiated	Adult-run	Adult-run

Mehan, 1979; Sinclair & Coulthard, 1975). Since it helps students to learn every kind of subject matter, researchers have encouraged teachers to bring more metacognitive talk into the classroom. Indeed, teachers who make more use of metacognitive talk produce students who are better able to understand their own and others' beliefs (Olson & Astington, 1993). The instructional importance of mentalistic talk is related not only to the understanding of beliefs but also to self-regulation (Paris & Paris, 2001). Metacognitive talk is commonly studied as a part of IRF sequences. Within a typical school classroom, these sequences are regularly initiated by a teacher question, followed by a student answer, and commonly closed by a teacher follow-up. The research presented below focuses on IRF sequences present at language lessons.

Lesson Structure

Lesson structure is the manner in which teachers partition their lessons as a set of more or less related events and activities. A focus on lesson structure is commonly related to the study of classroom management strategies. Still, successful strategies of this kind are not related to particular instructional approaches (Brophy, 2000). To relate patterns of time allocation to folk pedagogies, then, it is necessary to look for a framework other than class management. International comparative studies on teaching, such as the TIMSS studies (Givvin et al., 2005), provide the kind of framework that is needed and that has inspired my work.

FROM FOLK PEDAGOGY TO INSTRUCTIONAL PATTERN

Lesson talk and lesson structure are deeply informative of the instructional pattern prevailing within a particular educational system. Consequently researchers commonly look at these dimensions of teaching when they want to elucidate the cultural signature of education in a specific nation.

The connection between instructional pattern and folk pedagogy goes as follows. Patterns of teaching are based on shared mental models of teaching, or scripts related to the lesson's flow (Givvin et al., 2005). In turn, these scripts are anchored in the shared norms and values that shape teachers' folk pedagogies. These scripts mark how lessons are commonly structured and the discourse practices commonly adopted by teachers. Teachers create their lessons by spontaneously adopting the cultural script (or dominant folk pedagogy) as a guide.

The hypothesis that there are country-specific patterns of teaching has recently been challenged by an alternative hypothesis claiming that teacher practices are marked by institutional isomorphism between countries, so that the core teaching practices and beliefs are similar across the world (LeTendre, Baker, Akiba, Goesling, & Wiseman, 2001). According to this hypothesis, because of the way school as an institution has permeated diverse nations and because of the forces of globalization, which promote uniformity, school practices have become increasingly similar around the world (LeTendre et al., 2001). Whereas data supporting the global convergence hypothesis are based on self-reports, both the original TIMSS video studies and subsequent examination of their data have provided abundant evidence favoring the notion that country-specific teaching patterns can be described (Givvin et al., 2005; Hiebert & Stigler, 2000; Hiebert et al., 2005; Stigler, Gallimore, & Hiebert, 2000; Stigler & Hiebert, 1997, 2004; Stigler & National Center for Education Statistics, 1999).

The Chilean Setting

Where Chile is concerned, large-scale assessment of teaching has been implemented by two government initiatives: the *Docente Más* system (Teacher Plus, hereafter DM) and the *Asignación de Excelencia Pedagógica* system (Acknowledgment of Pedagogical Excellence, hereafter AEP). Whereas DM measures good-quality teaching in public schools by means of a compulsory assessment, AEP does so both in public and in partially subsidized private schools by means of a voluntary assessment (Manzi & Rosetti, 2004). Low returns on 15 years of investment in the large-scale side of education, measured by poor student achievement, drove policy makers in the Chilean government during recent years to set up these two large-scale systems of teaching assessment, whose goal is to assess the quality of teaching in Chilean schools.

These teaching assessment initiatives are part of the global trend to put teaching center stage in assessing educational quality. Indeed, both AEP and DM are aligned to a conceptual framework recently established by the Ministry of Education to set teaching standards: *El Marco para la Buena Enseñanza* (Framework for Good Teaching, hereafter MBE), which is inspired on Danielson's (1996) framework for teaching. In particular, the MBE endorses the following standards:

1. Provisions for learning: Teachers need a deep knowledge of the subject matter and of the pedagogical aspects required for a successful learning process.
2. A positive learning environment: Teachers must develop an ability to create a learning environment where all students feel trusted, accepted and respected.
3. Learner-centered teaching: Teachers must develop an ability to implement an educational setting that fosters students' commitment to the learning process.
4. Professional responsibility: A teacher must develop an ability to reflect upon his or her teaching practices and an ability to establish positive relationships with colleagues, supervisors, parents, and the school community (Ministerio de Educación, 2003).

The MBE provides a set of criteria for each one of its standards. In turn, these criteria are described through a number of more particular definitions. For instance, the standard "Provisions for teaching" includes the criterion "Mastering the principles and concepts of the disciplines taught by the teacher," which, in turn, includes five elements, one being "Mastering new perspectives and new developments in the disciplines taught by the teacher" (Ministerio de Educación, 2003). By aligning DM with the MBE, its implementation looks for a renewal of teaching habits. Whether the implementation of the MBE runs against the prevalent adult-run model is still an open question.

Since our data involve a reanalysis of films collected in the context of the DM initiative, a description of DM is in order. DM implementation started in the year 2003. As the Chilean government scheduled a gradual implementation of the system, each year new teachers have joined the assessment process. Thus, during the last 5 years of implementation, DM has evaluated more than fifty thousand teachers. To perform the assessment, DM uses four instruments: a self-assessment, an interview made by a peer, reports by third parties, and a portfolio, which includes, first, a video survey of a class conducted by the teacher and, second, teacher planning of three lessons. Based on the scores achieved by the teachers on the four instruments, DM differentiates teachers in four performance groups: outstanding, proficient, basic, and unsatisfactory (Manzi, 2007).

The illustrations we discuss below reanalyzed the video-survey component of the portfolio teachers produced for the 2003 and 2005 rounds of assessment. It is worth noting that the distribution of teachers across the four levels of competence remained quite stable in the following rounds of appraisal (Manzi et al., 2008), although the last evaluation saw an increase of teachers assessed as proficient.

Next, I proceed to illustrate three components of different studies, both finished and ongoing, which look to elucidate what the cultural markers are in Chilean instructional patterns.

THREE CHILEAN ILLUSTRATIONS

Illustrations 1 and 2: Communication Patterns in Elementary and Middle-School Teaching of Language

Next, I present the coding scheme and main results obtained in two studies intending to assess the communication patterns present in teacher talk of elementary and middle-school teaching of language. Results for elementary school have been reported elsewhere (Preiss, 2005, 2009) so we summarize here only those aspects that are relevant for the sake of the argument. Middle-school results are as yet unpublished. Here, we partially report those that allow one to complete the picture advanced by the elementary school study. A complete description of these results is forthcoming.

Coding Scheme

Save by small variations, both studies share the same coding scheme. The coding scheme for Teacher Talk (TT, hereafter) included two broad categories: teacher questions and teacher follow-ups. Teacher questions are all the questions the teacher asked the whole class or individual students publicly. Teacher follow-ups are all the public responses made by the teacher to students' responses. All of these codes were adjusted from Wells (1999) and from the TIMSS 1995 video study (Stigler et al., 2000). I selected and merged those codes that were more relevant to the Chilean context.

Teacher questions were coded as follows:

1. Controls: The purpose of these questions is to regulate the lesson flow. The teacher assesses whether students are keeping track of the lesson and organizes the allocation of tasks.

2. Inform. The purpose of these questions is to check for information. The teacher assesses whether students know some information or are following the information the teacher is delivering in the class.
3. Implement. The purpose of these questions is to make students use a verbal skill. In particular, the teacher trains the students in literate language.
4. Elaborate. The purpose of these questions is to ask the students to reflect on their previous work and express their beliefs. The teacher provides students with opportunities to think about what they have done or said, to collaborate with the class, and to negotiate the meaning of their work. Usually, these questions involve questions that ask Why? What for?

In addition, the middle school study applied the following codes:

5. Demand. Teacher demands a physical action.
6. Opinion. Teacher asks students' opinions.
7. Review. Teacher asks a question that intends to review contents previously taught.

Teachers' follow-ups were coded using the following codes:

1. Monosyllable/Neutral. The teacher responds to students' replies with a monosyllabic word, such as "uh uh" or "OK," or with an answer with no instructional consequences.
2. Repeat. The teacher answers to students' contributions with a partial or total repetition of the students' responses.
3. Evaluate. The teacher makes an explicit assessment of the student responses. In addition, this code included all those follow-ups that had an evaluative intention, such as "Are you sure of your answer?"
4. Reformulate. The teacher capitalizes on a student contribution to explore the content.

In addition, for middle school, we coded the kind of lecturing that prevailed in between IRF sequences. Codes used were task-related talk for talk associated with teacher management of instructional activities; content talk for talk associated with subject matter; and metacognitive talk for talk involving mentalistic references.

Procedure

Coding of TT focused on the lessons' beginning; that is, the first 10 minutes of each lesson. A focus on the first 10 minutes of the class allowed the study to code an uninterrupted segment of TT, whereas a random selection of segments might have resulted in the selection of significant amounts of lesson time without public talk. A transcription of the lesson was used for definitive coding.

Two independent coders coded the transcripts. Coders were instructed to code as follow-ups all teacher replies to students' responses that resulted from a teacher's question. All kinds of student responses were considered—namely, individual answers, group answers, or whole-class answers. If there was no teacher question anticipating a student response, however, no follow-up was identified. Replies to students that did not result from teachers' questions, such as directions or commands, were not considered. Moreover, TT that was not related to instructional content or involved disciplining students was not considered.

After coders reached agreement, each coder assessed part of the pending material. When data from both were available, a coder was randomly selected to provide the data for final data analysis.

Main Results

Elementary School Study From the universe of 3,740 elementary school teachers assessed by DM in 2003, material on 128 teachers was sampled for this study. Consistency was estimated using Pearson's R and all of the consistency values were acceptable ($r > .8$). Consensus was estimated by the index of percent agreement using adjacent categories. Consensus values were acceptable for *Implement, Elaborate, Monosyllable/Neutral* and *Reformulate* ($p > .7$) and moderate for *Inform* and *Evalua*te ($p > .6$). Because the focus of the coding process for TT was on consistency and the consensus data did not threaten the validity of my results, all of the variables were retained for analysis.

Observation of means and standard deviations for Questions and Follow-ups shows that there is a sizable amount of variation in the number of questions and follow-ups teachers produce. During the first 10 minutes of a lesson, teachers on average asked 26.53 questions (SD = 14.88) and there were 18.60 follow-ups (SD = 11.77). The teacher who posed the most questions asked 61 questions of the class, whereas the teacher who introduced the least number of questions asked just one. Similarly, teachers posed a maximum of 59 and a minimum of zero follow-up questions.

It was my expectation that by summarizing the correlations between different kinds of questions and follow-ups, a principal components analysis would clarify the dominant communication patterns in the lessons studied. Consequently, an exploratory principal components analysis with varimax rotation was run through SPSS on the eight kinds of TT variables for a sample of 128 teachers. Eigenvalues for the first three components were all larger than 1 and—after the fourth component—changes in successive Eigenvalues were small. The three-component solution explained 65.80% of the total variance. Extraction communality values ranged from .556 to .735. *Inform, Evaluate* and *Reformulate* had communality values marginally close to .6, so they were only moderately defined by the solution. All of the other variables had communality values over .68 and consequently were well defined by the solution.

With a cutoff of .5 for the inclusion of a variable in the interpretation of a component, all of the variables loaded on a component, and there were no complex variables. Loadings of variables on components are shown in Table 12.2. Each of these components depicts a communicative cycle. The first component involves *Implement* and *Control* questions and *Evaluate* follow-ups. This component, then, demonstrates a *procedural communicative cycle.* The second component involves *Inform* questions and *Reformulate* and *Repeat* follow-ups. This component we will call a *declarative communicative cycle.* The third component involves *Elaborate* questions and *Monosyllable/neutral* follow-ups. This component depicts what I have called an *incomplete internalist communicative cycle.*

In summary, three communicative patterns were unveiled. Two of these patterns match the two forms adopted by an externalist folk pedagogy. They show that Chilean elementary school teachers can focus their talk either on the systematic training of skills or on the delivery of information. The third pattern is reflective of teachers' lack of ability to fully implement an internalist pedagogy. When asking questions that might tap metacognitive skills or students' personal reasoning, teachers do not pursue students' contributions further.

Middle-School Study From the universe of 714 middle-school language teachers assessed by DM in 2005, material on 117 participants was sampled for further analysis. As before, consistency was estimated using Pearson's R and consensus was estimated by the index of percent

Table 12.2

ELEMENTARY SCHOOL LANGUAGE LESSON STUDY ROTATED COMPONENT MATRIX[a] FOR QUESTIONS AND FOLLOW-UPS

	COMPONENT		
	1	2	3
Implement	.839	.050	−.123
Evaluate	.757	.095	.067
Control	.626	.451	.292
Inform	.206	.741	.076
Reformulate	−.085	.740	−.016
Repeat	.406	.703	.275
Elaborate	−.207	.154	.801
Monosyllabic	.250	.018	.786

[a]Rotation converged in 5 iterations.
Notes: Extraction method: principal component analysis. Rotation method: Varimax with Kaiser normalization.

agreement using adjacent categories. All of the consistency values were very good ($r > .86$), but the code for elaborate was acceptable ($r > .78$) and the code for metacognitive talk was moderate ($r > .66$). Consensus values were acceptable for all the codes ($p > .83$). Based on these results, I decided to retain all the variables for subsequent analysis.

Observation of means and standard deviations for *Questions* and *Follow-ups* shows very similar results to those of the elementary school study. During the first 10 minutes of a lesson, teachers on average asked 24.16 questions ($SD = 11.47$) and 20.97 follow-ups ($SD = 12.48$). The teacher who posed the most questions asked 53 questions of the class, whereas the teacher who introduced the least number of questions asked just 4. Similarly, teachers posed a maximum of 73 and a minimum of zero follow-ups.

Inspection of the means of the variables revealed that the means for *Opinion* and *Metacognitive Talk* were close to zero, so these variables were dropped of subsequent analysis. Then, again an exploratory principal-components analysis with varimax rotation was run through

SPSS on the 12 kinds of TT variables for a sample of 117 teachers. Eigenvalues for the first four components were all larger than 1. The solution explained 60.93% of the total variance. Extraction communality values ranged from .4 to .76. *Elaborate* and *Content Talk* had communality values close to 4, so they were not well defined by the solution. Indeed, *Elaborate* did not load in any component. *Demand, Implement, Task-related talk,* and *Evaluate* had communality values between .5 and .6, so they were only moderately defined by the solution. All of the other variables had communality values over .66 and consequently were well defined by the solution. With a cutoff of .5 for the inclusion of a variable in the interpretation of a component, all of the variables but *Elaborate* loaded on a component, and there were no complex variables. Loadings of variables on components are shown in Table 12.3.

Table 12.3

MIDDLE SCHOOL LANGUAGE LESSON STUDY ROTATED COMPONENT MATRIX FOR QUESTIONS AND FOLLOW-UPS[a]

	COMPONENT			
	1	2	3	4
Inform	.787	.113	.100	−.213
Reformulate	.735	−.186	.026	.348
Content talk	.606	−.040	−.168	−.052
Elaborate	.495	.212	.080	.344
Evaluate	.152	.748	.064	.042
Demand	−.185	.721	.014	−.150
Implement	.104	.621	−.123	.329
Control	.037	−.079	.857	.131
Monosyllabic	.043	−.043	.788	−.199
Task-related talk	−.213	.350	.549	.274
Review	−.140	−.053	−.041	.858
Repeat	.417	.239	.155	.666

[a]Rotation converged in 6 iterations.
Notes: Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Each of these components depicts a communicative cycle. The first component involves *Content talk*, *Inform* questions, and *Reformulate* follow-ups. This component, then, demonstrates a *content-based communicative cycle.* The second component involves *Implement* and *Demand* questions and *Evaluate* follow-ups. This component we will call a *skill-drilling communicative cycle*. The third component involves *Task-focused* talk, *Control* questions and *Monosyllable/Neutral* follow-ups. This component depicts a *class-management communicative cycle.* Last but not least, a fourth component involves *Review* questions and *Repeat* follow-ups and demonstrates a communicative cycle focused on the *review of content* previously taught.

In summary, four communicative patterns were unveiled. The first component suggests that teachers engage in some form of metacognitive talk when delivering new content, presumably declarative content. Two of these patterns are related to the repeated practice of skills, either focused on implementing some form of basic language training or on monitoring task completion. Finally, when teachers engage students in reviewing content previously taught, they do not explore that content further.

Illustration 3: Lesson Structure in Middle-School Teaching of Mathematics

Coding Scheme

Codes for this study were inspired by those of the TIMSS 1999 video study, so their definitions are close to those presented in this study's report. Owing to limitations of space, I provide only descriptive results for two dimensions studied: these are mathematical work and social organization of the lesson, which are both very informative of the instructional pattern evident in the lessons studied.

Social Organization of the Lesson

The coding scheme for social organization of the lesson was as follows.

1. Teacher-led public interaction: This code refers to a typical whole-class situation involving public talk. The teacher gives a presentation to the whole class in a situation of joint attention between

the teacher and the students. The activities implemented are manifold and may involve learning something new, reviewing something already seen, or problem-solving tasks. This code is specified by two subcodes:

a. Question-and-answer exchanges: Interaction is based on a sequence of, at least, three teacher-initiated question-and-answer exchanges. It includes those exchanges focused on shared assessment of problems previously presented by the teacher for the class to work on independently.
b. Lecturing: Teacher presents mathematical contents without initiating question-and-answer exchanges with the students.

2. Student-led public interaction. A student leads a public discussion or makes a presentation to his or her classmates.
3. Individual private work. The students perform individual seat work, whereas the teacher circulates around the classroom, publicly or privately interacting with groups or individual students. Although the teacher may talk aloud to attend students' questions, resolve doubts, or reinforce his or her instructions concerning the activity, attention is not shared between the teacher and the students.
4. Group private work. Students perform group seat work, whereas the teacher circulates around the classroom, publicly or privately interacting with groups or individual students. Attention is not shared between the teacher and the students but within groups of students.

Mathematical Work

The coding scheme for mathematical work was as follows:

1. Nonmathematical work: for instance, socialization, disciplining students, motivating students, and so on.
2. Copying content: 2 minutes or more copying content from the board.
3. Reviewing content: the teacher focuses on content previously taught without engaging in problem solving.
4. Introducing new content: the teacher presents new material to the students either by means of problems or by defining concepts.
5. Practicing content: the teacher focuses on content practice either by conventional problem solving or other means.

Codes for introducing new content and practicing content could be specified by the following subcodes:

a. Problem solving: Mathematical work involving problem solving (for instance, addition, subtraction, algebra, equations, measurement, geometry, graphing). As in the TIMSS 1999 video survey scheme, problem solving may involve:
   i. Independent problems—that is, problems that are presented and worked on an individual basis, either publicly or privately
   ii. Concurrent problems—that is, problems that are commonly taken from texts or guides to be worked on privately, although some of them might eventually be discussed in the full class
   iii. Answered-only problems—that is, problems that have been previously done during either tests or lesson activities

b. Other mathematical work: Mathematical work that does not involve problem solving and that may involve presenting new mathematical definitions or concepts, relating mathematical ideas to the real world, or play not involving problem solving

c. Mixed independent problem solving and other mathematical work: Mathematical work that involves a mixture of independent problem solving and contextual activities, such as those mentioned above

d. Mixed concurrent problem solving and other mathematical work: Mathematical work that involves a mixture of independent problem solving and contextual activities, such as those mentioned above

Procedure

We applied the codes presented below to all the behaviors that lasted 1 minute or more. If there was a behavior that lasted less than 1 minute, its time duration was added to the subsequent behavior. Two coders analyzed the lessons. Double coding was performed in approximately one-third of the videos. Consensus was estimated by the index of percent agreement, as reported by The Observer. Consensus values were acceptable for all the codes ($p > .75$). To solve disagreements, coders contrasted and discussed their particular results and reached a new agreement based on their joint observation of the material.

Main Results

As shown in Table 12.4, teacher-led public interaction was the main activity performed by teachers in the lessons observed. It took, on average, the largest amount of lesson time and was the most frequently implemented activity. The favored way for teachers to lead public interaction was by means of question-and-answer exchanges instead of lectures. When students were doing seat work, data suggest a preference for individual work instead of group work.

Inspection of means for mathematical work reveals that the activity favored by teachers is practicing content, followed by introducing content. In comparison to these two activities, time allocated to reviewing content was relatively minor (see Table 12.5).

Because of teacher emphasis on content practice, I decided to determine what the teachers' favored kind of practice was. As shown in Table 12.6 and Figure 12.1, teachers had a preference for practicing concurrent problems, although some form of mix between concurrent problem solving and other mathematical work is present. Still, the use of guides or problem series seems to be a cultural signature of the observed lessons. Indeed, inspection of box plots for the different kinds of practice reveals that those teachers engaging in a form of practice other than solving concurrent problems are statistical outliers.

In summary, middle-school mathematics lessons possess a structure marked by teacher-led problem solving of concurrent problems. Most of the analyzed lessons were organized around the repeated practice of mathematical problems. Like teachers in the United States, Chilean teachers present students with definitions and procedures that they are expected to practice later on. Although Chilean teachers seem to adopt a dialogical approach that pervades conventional lecturing, most of students' independent work is done individually instead of in groups. Taking into consideration these cultural markers, I described these lessons as involving the *private appropriation of terms and procedures.*

CONCLUSION

The three illustrations above show that most of the Chilean teachers observed are aligned with externalist pedagogies. The data also dem-

Table 12.4

SOCIAL ORGANIZATION OF THE LESSON IN CHILEAN MIDDLE SCHOOL MATHEMATICS LESSONS (N = 114)[a]

	TOTAL DURATION		NUMBER OF SEGMENTS		MEAN PER SEGMENT	
ACTIVITY	*M*	*SD*	*M*	*SD*	*M*	*SD*
Teacher-led public interaction Question-and-answer exchanges	1103.34	512.27	3.14	1.32	412.31	325.85
Teacher-led public interaction Lecturing	208.45	228.56	1.42	1.17	112.02	106.30
Private group work	422.44	569.54	.6228	.83497	373.5439	533.74742
Private individual work	569.51	562.45	1.3860	1.28648	354.9386	438.28821

[a]Variables are measured in seconds.

Table 12.5

MATHEMATICAL WORK IN CHILEAN MIDDLE SCHOOL MATHEMATICS LESSONS

ACTIVITY	NUMBER OF SEGMENTS *M*	NUMBER OF SEGMENTS *SD*	SECONDS PER SEGMENT *M*	SECONDS PER SEGMENT *SD*
Introducing content	1.2895	1.11088	183.9737	207.39089
Practicing content	2.5702	1.31680	880.9211	554.73296
Reviewing content	.5088	.6274	60.3947	88.21761

Valid *N* (listwise) 114.

Table 12.6

NUMBER OF SEGMENTS FOR MATHEMATICAL WORK VARIABLES ($N = 114$)

BEHAVIOR	*M*	*SD*
Practicing independent problems	.0965	.39843
Practicing concurrent problems	1.5263	1.17654
Practicing answered-only problems	.0088	.09366
Practicing other mathematical work	.2719	.56888
Practicing independent mixed problem solving and other mathematical work	.0175	.13187
Practicing mixed concurrent problem solving and other mathematical work	.6404	.74224

onstrate the main weaknesses faced by Chilean schoolteachers: for language lessons, a conversational style that does not capitalize on students metacognition; for mathematics lessons, a uniform repetitive lesson structure and an orientation to a narrow set of skills that does not match the potential diversity of students.

A majority of Chilean language teachers address their students by means of objective and informative talk. Both for elementary and

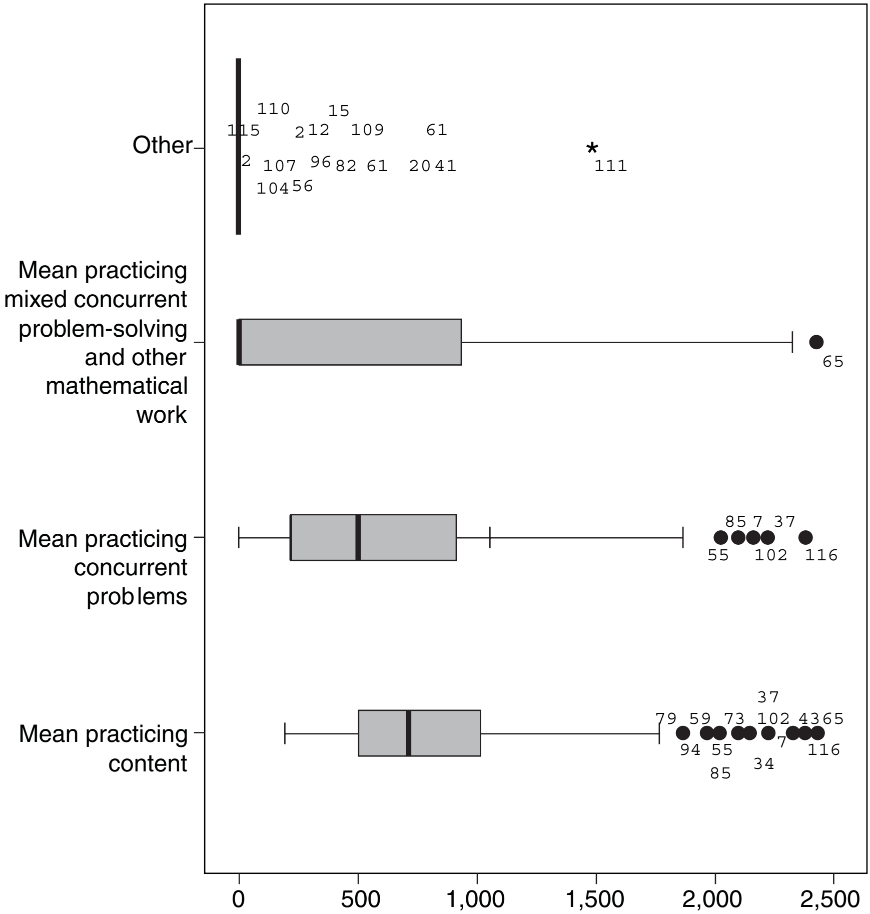


Figure 12.1 Practice kinds in middle school mathematics lessons.

middle-school teaching of language, teachers focus their talk either on the systematic training of skills or on the delivery of information. Data for mathematics teachers replicated findings I reported elsewhere for elementary school language teachers: specifically, Chilean schoolteachers' proclivity to produce class work for a large proportion of the lesson time (Preiss, 2005, 2009). Chilean mathematics lessons engaged the teacher working with the class as a whole more than 50% of the time. Summaries of existing research in the United States indicated that 65% of the lesson is commonly devoted to seat work (Gallego, Cole, & The Laboratory of Comparative Human Cognition, 2001).

In a similar way as for elementary school language teachers, middle-school mathematics teachers have a preference for leading short communicative exchanges back and forth with students. In each case, students play a marginal public role, either individually or in groups. When students are performing private work, most of the time they are working individually. The data suggest that, as in the United States, teachers present students with definitions and procedures, which they have to memorize and practice later on. While Stigler and Hiebert described American lessons as "learning terms and practicing procedures" (1999, p. 27), I described Chilean lessons as "private appropriation of terms and procedures" because of the Chilean emphasis on individual work. The emphasis of Chilean teachers on practicing concurrent problems may indicate an overemphasis on skill drilling instead of mathematical understanding. Additionally, it may suggest that mathematical lessons possess a very simple structure and lack a structured presentation of mathematical reasoning. Recent data from new groups of teachers assessed by DM have shown that lack of structure is one of the main weaknesses of the lessons assessed (Manzi, 2007). Consequently, this is one of the areas calling for further research.

These data complement others showing that Chile must confront issues of teacher preparation and classroom reform. Recently a voluntary diagnostic assessment taken by 1994 graduates trained in 80% of all teacher schools in Chile revealed that teacher preparation faces serious challenges, particularly in the mastery of the subject matter teachers are expected to teach (Ministerio de Educación, 2009). Future research should explore whether the prevailing folk pedagogies are related to the issues confronted by teacher training and teacher recruitment. Specifically, one of the most interesting and pressing follow-up questions to the research presented here concerns the interaction of teachers' mastery of subject matter with teaching pedagogies.

As regards educational reform initiatives, owing to issues of cultural compatibility raised earlier, it is not clear that recommendations from educational theories originating outside the Chilean context can be adopted without modification. However, any effort of educational reform must meet contemporary educational advances and look for an elective affinity with those theories that, produced outside of Chile, are more suitable to working with the peculiarities of the Chilean imagination. The current emphasis of some educators on dialogical inquiry (Wells, 1999) and lesson study (Stepanek, Appel, Leong, Mangan, & Mitchell, 2007), may match the dispositions of current Chilean culture and the current capabilities of the Chilean educational system nicely.

In particular, the framework of dialogical inquiry may allow teachers to develop their own metacognitive models—and their own implicit pedagogies—about their teaching competence, where they can use a frame of reference that capitalizes on the propensity to dialogue observed here. The opportunity for teachers to inspect their own classes and attend to the kind of questions and follow-ups they produce would give teachers a chance to think about the kinds of interaction they are promoting within the class, the pedagogical effect of these interactions, and whether they meet their own pedagogical goals. In fact, the same material used in DM for assessment can be used for pedagogical purposes, adopting the lesson study approach. The fact that all the material collected by DM is stored by the main state center for teacher development—the Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (Center for Improvement, Experimentation and Pedagogical Research) suggests that the conditions are ideal for a professional development initiative of this kind, whereby teachers' implicit pedagogies could be challenged, nurtured, and developed by peers. An initiative such as this would demonstrate a serious commitment to the development of expertise in teachers and recognition of the social importance of this expertise.

Last but not least, my results provide clear support to the hypothesis that there are country-specific patterns of teaching (Givvin et al., 2005). Although the relation between cultural patterns of teaching and educational achievements is still controversial, future public policy initiatives for educational reform should take into consideration the cultural basis of the teaching patterns they want to transform and how permeable these patterns are to initiatives originated outside of the classrooms. Consequently, initiatives that focus on the study of lesson practices, either by using DM database or by implementing local initiatives at school, may be the most effective way to address educational reform in Chile.

REFERENCES

Alexander, R. (2000). *Culture and pedagogy. International comparisons in primary education.* London: Blackwell.

Alexander, R. (2001). Border crossings: Towards a comparative pedagogy. *Comparative Education, 37*(4), 507–523.

Avalos, B. (2004). Teacher regulatory forces and accountability policies in Chile: From public servants to accountable professionals. *Research Papers in Education, 19*(1), 67–85.

Barber, M., & Mourshed, M. (2007). *How the world's best-performing school systems come out on top.* London: McKinsey.

Blair, C., Gamson, D., Thorne, S., & Baker, D. (2005). Rising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. *Intelligence, 33*(1), 93–106.

Bransford, J., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Brophy, J. (2000). Classrooms: Processes and management. In A. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 2, pp. 95–102). Washington, DC: American Psychological Association.

Bruner, J. S. (1990). Acts of meaning (p. 179). *Cambridge, MA: Harvard University Press.*

Bruner, J. S. (1996). *The culture of education.* Cambridge, MA: Harvard University Press.

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.

Clark, C. M., & Peterson, P. L. (1986). Teachers thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 255–296). New York: Macmillan.

Cole, M. (2005). Cross-cultural and historical perspectives on the developmental consequences of education. *Human Development, 48*(4), 195–216.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Dinsmore, D., Alexander, P., & Loughlin, S. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20*(4), 391–409.

Fenstermacher, G. D., & Richardson, V. (2005). On making determinations of quality in teaching. *Teachers College Record, 107*(1), 186–213.

Givvin, K. B., Hiebert, J., Jacobs, J., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review, 49*(3), 311–343.

Gorard, S. (2001). International comparisons of school effectiveness: The second component of the "crisis account" in England? *Comparative Education, 37*(3), 279–296.

Haye, A., & Pacheco, V. (1995). El papel de las representaciones sociales en la crisis de la educacion chilena [The role of social representations in the crisis of Chilean education]. *Psykhe, 4*(1), 25–38.

Healy, T., Côté, S., Helliwell, J. F., Field, S., Centre for Educational Research and Innovation, & Organisation for Economic Co-operation and Development. (2001). *The well-being of nations: The role of human and social capital.* Paris: Organisation for Economic Cooperation and Development.

Hiebert, J., & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS video study. *Elementary School Journal, 101*(1), 3–20.

Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H., Smith, M., et al. (2005). Mathematics teaching in the United States today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis, 27*(2), 111–132.

LeTendre, G., Baker, D., Akiba, M., Goesling, B., & Wiseman, A. (2001). Teachers' work: Institutional isomorphism and cultural variation in the U.S., Germany and Japan. *Educational Researcher, 30*(6), 3–15.

Loera, A. (2006). *La práctica pedagógica videograbada. [Teaching practices videotaped].* México, DF: UPN.

Manzi, J. (2007). Evaluación docente: Antecedentes, resultados y proyecciones [electronic version]. *Seminario evaluación docente en Chile: Fundamentos, experiencias y resultados.* Retrieved January 2, 2008, from http://mideuc.cl/seminario/presentaciones.html

Manzi, J., & Rosetti, J. (2004). La evaluación educacional en Chile: Antecedentes y desarrollo reciente. [Educational assessment in Chile: Antecedents and recent developments]. *Persona y Sociedad, 18*(3), 221–243.

Mayer, R. E. (1998). Cognitive theory for education: What teachers need to know. In N. M. Lambert & B. L. McCombs (Eds.), *How students learn: Reforming schools through learner-centered education* (pp. 353–377). Washington, DC: American Psychological Association.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.

Ministerio de Educación. (2003). *Marco para la buena enseñanza [Framework for good quality teaching].* Santiago, Chile: CPEIPo.

Ministerio de Educación. (2009). Presentación de resultados. Evaluación diagnóstica inicia 2008 [Results presentation. Diagnosis assesment inicia 2008] [electronic version]. Retrieved April 15th, 2009, from http://www.cpeip.cl/recursos/200904071657290.presentacion_resultados_inicia09.pdf

Morandé, P. (1984). *Cultura y modernización en América Latina: ensayo sociológico acerca de la crisis del desarrollismo y de su superación. [Culture and modernity in Latin America: sociological essay about the crisis of the theories of social development and overcoming them].* Santiago: Instituto de Sociología, Pontificia Universidad Católica de Chile.

Nuñez, I. (2002). La formación de docentes. Notas históricas [Teacher training. Historical notes]. In B. Avalos (Ed.), *Profesores para Chile, historia de un proyecto [Teachers for Chile, history of a project]* (pp. 14–39). Santiago: Ministerio de Educación.

Olson, D. R. (1994). *The world on paper: The conceptual and cognitive implications of writing and reading.* New York: Cambridge University Press.

Olson, D. R. (2003). *Psychological theory and educational reform: How school remakes mind and society.* Cambridge, UK: Cambridge University Press.

Olson, D. R. (2005). Technology and intelligence in a literate society. In R. J. Sternberg & D. Preiss (Eds.), *Intelligence and technology. The impact of tools on the nature and development of human abilities* (pp. 55–68). Mahwah, NJ: Erlbaum.

Olson, D. R., & Astington, J. W. (1993). Thinking about thinking: Learning how to take statements and hold beliefs. *Educational Psychologist, 28*(1), 7–23.

Olson, D. R., & Bruner, J. S. (1996). Folk psychology and folk pedagogy. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching, and schooling* (pp. 9–27). Cambridge, MA: Blackwell.

Olson, D. R., & Torrance, N. (Eds.). (1996). *The handbook of education and human development: New models of learning, teaching, and schooling.* Cambridge, MA: Blackwell.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36*(2), 89–101.

Paz, O. (1985). *The labyrinth of solitude: Life and thought in Mexico.* New York: Viking Penguin.

Preiss, D. D. (2005). *The task of teaching: Folk pedagogy and instructional practice in Chile's elementary education.* Unpublished Ph.D. dissertation, New Haven, CT: Yale University.

Preiss, D. D. (2009). The Chilean instructional pattern for the teaching of language: A video-survey study based on a national program for the assessment of teaching. *Learning and Individual Differences, 19,* 1–11.

Preiss, D. D., & Strasser, K. (2006). Creativity in Latin America: Views from psychology, humanities, and the arts. In J. C. Kaufman & R. J. Sternberg (Eds.), *The international handbook of creativity* (pp. 39–67). New York: Cambridge University Press.

Rogoff, B., Matusov, E., & White, C. (1996). Models of teaching and learning: Participation in a community of learners. In N. Torrance & D. R. Olson (Eds.), *The handbook of education and human development* (pp. 388–414). Cambridge, MA: Blackwell.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15,* 4–14.

Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.

Stepanek, J., Appel, G., Leong, M., Mangan, M. T., & Mitchell, M. (2007). *Leading lesson study: A practical guide for teachers and facilitators.* Thousand Oaks, CA: Corwin Press.

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 & Social Psychology, 49*(3), 607–627.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence.* New York: Cambridge University Press.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., et al. (2000). *Practical intelligence in everyday life.* New York: Cambridge University Press.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist, 35,* 87–100.

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction—An overview of TIMSS video study. *Phi Delta Kappan, 79*(1), 14–21.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap. Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 12–17.

Stigler, J. W., & National Center for Education Statistics. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education, 44*(4), 387–407.

Takayama, K., & Apple, M. W. (2008). The cultural politics of borrowing: Japan, Britain, and the narrative of educational crisis. *British Journal of Sociology of Education, 29*(3), 289–301.

Torff, B., & Sternberg, R. J. (2001). *Understanding and teaching the intuitive mind: Student and teacher learning.* Mahwah, NJ: Erlbaum.

Weber, M. (2001). *The Protestant ethic and the spirit of capitalism.* London: Routledge.

Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education.* New York: Cambridge University Press.

This page intentionally left blank

PART V

Psychological Perspectives on Educational Reform

This page intentionally left blank

13

A New Role for Schools: Providing Child Care and Family Support Services

EDWARD ZIGLER AND MATIA FINN-STEVENSON

Recent studies in brain research underscore the importance of the early years of life and have prompted widespread interest in addressing the educational needs of children before they begin formal schooling. Approximately 40 states and the federal government have invested in preschool programs and schools around the country; these are opening their doors to increasingly younger children, providing not only care and education for preschoolers but also other family support programs beginning at the conception of the child and throughout the school-age years. In this chapter we describe the School of the 21st Century (21C), a program that has been a model for this approach. 21C has been implemented in more than 1,300 schools, enabling educators to address the needs of children and families. It has also paved the way for other school-based and school-linked services that have been implemented in recent years. In this chapter we describe 21C and review its research base, implementation, and impact, showing how the program is transforming education in America.

BACKGROUND

What Is the School of the 21st Century?

The School of the 21st Century is often referred to as "a program," but it is actually made up of several programs and services provided in the school from conception to age 12. At the core of 21C are child-care programs. First, there are all-day, year-'round child care for children aged 3 and 4 prior to their entry to kindergarten. We refer here to developmentally appropriate care which ensures that children have opportunities for play and social interaction as well as preschool educational activities so that they can acquire the skills needed for later learning. Second, there is before- and after-school as well as summer vacation care for children from kindergarten to age 12. The focus of many similar programs for children of school age is on providing enrichment activities that are academic in focus. In 21C, the emphasis is on providing children with the opportunity to choose among various types of activities, including recreational and enrichment opportunities. School-age child care in 21C is not limited to academic enrichment and homework. The rationale here is that for optimal development, children need time off from academic tasks and can profit from a wide array of youth development activities.

Also included in 21C are several outreach components. One is home visitation to families with newborns and young children, from conception to age 3, patterned after the Parents As Teachers (PAT) program. Parent educators visit the home, offer information to parents about their children's development, and conduct screenings for any potential developmental and learning problems. This provides opportunities for referral to early intervention services, before problems become exacerbated. 21C also includes outreach to family child care and other child-care providers in the community; information and referral for various services families may need; health, mental health, and nutrition education and services; and other services that may be needed in different communities. Although these 21C components are described separately, they are conceptualized as part of the 21C "umbrella" and are coordinated as a whole.

Need and Rationale for the Program

Impact of Societal Changes on Children's Development

Driving the development of 21C has been the continued need for child and family support services brought about by the societal changes that have occurred over the past several decades: changes in the family structure, including divorce and increases in the number of children living in single parent families; high mobility, especially in families with young children; and what some have come to term as lack of social capital, or "social poverty," referring to the dearth of adults in the lives of children and weak ties between families and their neighbors and kin (Putnam, 1995). The result has been increased isolation and alienation, which leave parents to raise children without any help and support.

These societal changes, as well as the fact that many children live in poverty, mean that many children are growing up under stressful conditions. Often associated with these are child maltreatment and dysfunction in family life, which can have profound developmental and educational consequences, resulting not only in low academic achievement but also juvenile delinquency and other social problems (Duncan & Brooks-Gunn, 1997; Shonkoff & Phillips, 2000; Vinson, Baldry, & Hargreaves, 1996). The stress in children's lives often stems from several sources rather than a single one. For example, children whose parents divorce may live in a single-parent family and also experience periods of poverty. The more stress factors children encounter, the greater the likelihood that they will suffer developmental consequences (Rutter, 1980), although children who have access to some form of support show resilience in the face of adversity (Garmezy, 1985; Rutter, 1979).

The Early Years and School Readiness

The children likely to be affected by difficult life conditions are very young. The early years—from the prenatal stage and the birth of the child to about age 5—represent an important developmental period during which children are vulnerable to risks but also receptive to interventions. These points are underscored in a recent review of the research by Shonkoff and Phillips (2002), who note that a number of core concepts frame our understanding of the nature of human development, including these:

1. That parents' behaviors and choices play a major role in determining the growth trajectory of the child.
2. That human development is shaped by the ongoing interplay among sources of vulnerability and sources of resilience.
3. That the timing of early experiences can matter but, more often than not, the developing child remains vulnerable to risks and

open to protective influences throughout the early years of life and into adulthood.

4. That the course of development can be altered in early childhood by effective interventions that change the balance between risk and protection, thereby changing the odds in favor of more adaptive outcomes (Shonkoff & Phillips, 2000, p. 3).

We have long known about the importance of interventions during the early years, as noted in the development of Head Start (Zigler & Valentine, 1979) and other early intervention programs. More recently, there has been renewed interest in the early years. Policy makers, among others, began to focus on school readiness, a concept that recognizes that children's experiences before they start kindergarten have profound implications for development and can determine whether they will succeed or fail in school. This policy interest in school readiness is noted in the 1990 National Education Goals, Goal One of which states that "all children in America will start school ready to learn" (National Education Goals Panel, 1991). Included in Goal One are several objectives:

> That all disadvantaged children will have access to high quality and developmentally appropriate preschool programs to help them prepare for school; that every parent in America will be the child's first teacher and will devote time each day to helping his or her preschool child learn, and that parents will have access to training and support they need to accomplish this; that children will receive nutrition and health care needed to arrive at school with healthy minds and bodies, and that the number of low birth weight babies will be significantly reduced through enhanced prenatal health systems. (National Education Goals Panel, 1991, p. 61)

The policy interest in school readiness has been strengthened in part by recent studies on the brain that support much of what we have known from the social sciences regarding the role of early experiences (Zigler, Finn-Stevenson, & Hall, 2002). However, there has been a limitation to this policy approach, which emphasizes literacy, numeracy, and cognition. The emphasis, which includes a focus on academics, effectively ignores the concept of the whole child and the fact that several developmental pathways besides cognition are critical for overall development and children's ability to succeed academically, particularly the social–emotional pathway. Also ignored is the notion that learning, especially in young children, occurs in numerous ways, including through play and social interactions and in a variety of contexts.

Besides the narrow view of development and learning, current policy attention to early education and school readiness is not backed by sufficient financial resources (Gilliam & Zigler, 2001). Hence policy makers at the federal level target low-income children, and even under the best circumstances, only a portion of eligible low-income children can be served, as is the case with Head Start. The same is true at the state level. Many more states are supporting early childhood programs, but these serve only a small percentage of the eligible low-income children (Barnett, Hustedt, Friedman, Boyd, & Ainsworth, 2008).

The Need for Child Care

When financial resources are scarce, the focus on poor children is understandable. However, as we described earlier in the chapter, there are circumstances beyond poverty that affect all families with children, one of these being the inability to find good-quality, affordable child care.

The lack of good-quality, affordable child care has reached crisis proportions, with detrimental consequences for children, many of whom are in poor-quality programs. The need for child care is evident in the statistics: 65% of mothers with children under age 6 and 78% of mothers with children between the ages of 6 and 13 are in the labor force; among mothers with infants under age 1, 59% are in the labor force or actively looking for work (U.S. Census Bureau, 2000). With their parents working, 13 million infants and preschool children—or 3 out of every 5 children—are in child care. Among school-age children, the situation is even worse, with an estimated 7 million children left at home alone while their parents are working. The child-care arrangements for the children vary, but increasingly parents are using child-care centers and smaller numbers of parents utilize family child care (where a woman takes care of several children in her own home), in-home caregivers, or care provided by a relative.

The need for child care was recognized in 1970 at the White House Conference on children. The participants noted the need for good-quality affordable child care as the number one need facing America's children and families. In 1971 we were on the brink of developing a child-care system but were stymied by the veto of the Comprehensive Child Development Act. Since 1971, with the exception of the inadequate Child Care and Development Fund, advocates have attempted in vain to create an interest at the federal level not only in providing support for establishing a much-needed child-care system but also in ensuring quality care through federal guidelines for child-care licensing regulations. This did not occur, however, leaving each state to develop its own licensing regulations. Although the intent in state child-care licensing is to provide for children's needs and protect them from harm, many states are failing to adequately address the basic health, safety, and developmental needs of children (LeMoines, Morgan, & Azer, 2003; Young, Marsland, & Zigler, 1997).

The failure at the state level to ensure even a basic level of safety for children in child care is noted in several national studies, reviewed elsewhere (Zigler, Marsland, & Lord, 2009). These studies point to the physical and developmental consequences associated with poor-quality child care, including health and safety risks (Consumer Product Safety Commission, 1999), delayed language and reading skills, and increases in aggression and behavior problems (Phillips, 1995; Peisner-Feinberg et al., 1999).

Children who are likely to receive the lowest quality of care are those from working-class families. Children of families in poverty receive better-quality care because they have access to child-care subsidies and programs that are publicly supported and thus more likely to be monitored on the basis of quality standards. Middle- and higher-income families can afford to purchase better-quality care. Still, the cost of care is so high that many families are precluded from enrolling their children in high-quality facilities.

There are numerous other issues related to child care besides the high cost of care and lack of availability and access to good-quality, affordable programs. Examples include lack of care for children whose parents work in nontraditional fields and different shifts, lack of child care in poor neighborhoods, and the fact that not much is known about issues related to care for children with disabilities (Shonkoff & Phillips, 2002). These and other issues cannot be explored within the scope of this chapter but are detailed in other of our publications (Finn-Stevenson & Zigler, 1999; Zigler et al., 2002; Zigler, Gilliam, & Jones, 2006).

Also, although we focus here on care for very young children, the need for good-quality, affordable, and accessible care extends to school-age children as well. Since school schedules and vacations do not coincide with work schedules, school-age children need to be in child care, but when it is unavailable or too expensive, they are left home alone. Studies have found that children who are left home alone or with their peers when school is out are at risk for a host of problems, including crime, substance abuse, and behavioral problems (Vandell, & Shumow, 1999; Fight Crime Invest in Kids, 1997). Two other points are of relevance (a) the consequences for children of being left unsupervised when school is out include low academic achievement and (b) children's academic performance may be enhanced by attending school-age child care. Children attending good quality school-age care programs have been found to have better academic work habits and fewer behavioral problems than those who are home alone when school is out (Miller, O'Connor, Stringnano, & Joshi, 1996).

21C: A RESPONSE TO THE NEED FOR CHILD CARE

The issues we raise, albeit briefly, point to the lack of child care and the fact that there is no universal system of child care in the United States. Rather, we have a collection of child care services that differ in quality and make up a patchwork nonsystems. It is not merely additional services and more money that are needed to address the problem but rather the establishment of a system, because the current nonsystem is difficult to access and improve. The School of the 21st Century seeks to establish a much-needed child-care system by joining with and making use of the already existing educational system—providing child care in the school.

Guiding Principles

21C was conceptualized on the basis of six guiding principles.

First, in order to address the needs of all children and ensure their optimal development, a child-care system must become a national priority and part of the very structure of our society, as is the case with education. Stable, reliable, good-quality care—vital to children's development and well-being as well as their education—must be a central element of such a national system. In using the term *good-quality care* we refer to care that is developmentally appropriate, as opposed to basic supervision or mere babysitting, and that gives children opportunities for play, learning, and social interactions with peers and adults.

Second, good-quality care should be accessible to every child regardless of ethnic or socioeconomic group in order to avoid a two-tier system wherein some children receive good-quality care and others do not. Like education, child care can be universally available only if it is primarily a state-based system. The federal government's role is both to subsidize the care of the most needy, as it currently does, as well as to support research, evaluation, and other efforts to enhance the system and ensure the provision of good-quality care.

Third, the child-care system in this country must be based on the optimal development of children and should emphasize a whole-child approach that places equal weight on all developmental pathways: social, emotional, physical, and cognitive. For purposes of research, social scientists often regard each developmental domain separately and have in the past given more weight to cognitive development. However, it has become widely recognized that in reality, all aspects of growth and development are interdependent; they occur simultaneously and should thus be given equal attention. This third principle not only acknowledges the child as a whole but also reflects the fact that while child care is often regarded as primarily a service for parents, it is first and foremost an environment where children spend a significant amount of time. As such, its quality affects their growth and development.

Fourth, parents and those who care for and educate children must work together; hence the importance of parental involvement, which has been shown to be essential for children's optimal development. The importance of parental involvement is noted not only for programs for preschool and younger children but also for children already in school.

The fifth principle calls for recognition, support, and decent pay for child-care providers, since they play a crucial role in the quality of care children receive. This principle also encompasses the need for ongoing provider training as well as appropriate pay upgrades.

Sixth, a national child-care system must be flexible and adaptable. Because every family has particular child-care needs, a universal system must be able to provide a range of choices for child care. Inherent in this principle is the recognition that there are differences not only among families but also among communities.

If such a child-care system is to be readily accessible in terms of cost as well as location to all families, the most efficient way to implement it would be to tap into the already existing educational system. Our country has a trillion-dollar investment in public school buildings, which are supported by tax dollars and used for only part of the day 9 months a year. By capitalizing on this investment and incorporating a child-care system into the already existing public school system, we would be able to increase the supply of child care as well as to ensure equitable and affordable good-quality care to all children.

Two issues must be emphasized with regard to the concept of universally accessible child care: First, *universal* is a term indicating that regardless of socioeconomic class, all families should have access to good-quality, affordable care. This does not mean that child care should be compulsory. Rather, it means that it should be accessible to those who need and want such care. Second, given limited public funds, we cannot expect support for child care for all children. Hence, in the School of the 21st Century, the provision of child care is based on parental fees for child care calibrated to parental income.

FROM THEORY TO PRACTICE

21C represents the authors' work in applied research. The field of applied research is important because it provides an opportunity to further our understanding of child development and learning while also focusing on issues and programs central to improving the well-being of children and families, thereby moving from research to practice. The guiding principles outlined above are based on 21C's theoretical framework; they spell out the steps to be undertaken to ensure the optimal development of children and their success in school. The principles also provide direction for educators to implement the program.

Implementation is a key aspect in moving from theory to practice and developing programs. In the context of developing 21C for national implementation, we have come to appreciate that any reform program must have three well-developed components: a theory of development and learning represented in the idea for the program; a theory of action that focuses on the local context and the conditions under which implementation of the idea is likely to succeed (Fullan, 2001); and a theory of action that focuses on external factors and the national context, which refers to what is needed if the program is to be implemented to scale (Berends, Kirby, Naftel, & Mckelvey, 2000; Kirby, Berends, & Naftel, 2001). In the following section we discuss the implementation of 21C within the context of these latter two aspects, focusing first on the local context.

Implementation of 21C in Schools

Recall from the previous pages that 21C includes several different programs and services: all-day, year-'round child care for preschoolers, child care for school-age children during times when school is out, and several other outreach and family support services, all of which are provided in the school. The use of the school is one of the unique aspects of 21C. However, 21C calls for more than simply utilizing space in the school building. Rather, the program is actually administered by and becomes part of the very fabric of the school.[1] Schools that implement 21C are changed substantially, opening their doors to young children and beginning to work with parents soon after the child's conception. The academic calendar is changed as well, with schools operating programs the entire day—from 6 in the morning to 6 in the evening—year 'round. One such school, in Bridgeport, Connecticut, is simply known as Six to Six. It is not only the provision of services that accounts for the change but, more importantly, the transformations that occur when educators take responsibility for addressing the various nonacademic needs children and families may have.

Factors Facilitating Implementation

As a whole-school reform program, 21C presents unique challenges to schools, since establishing a system for child care and working with young children are outside the traditional mission of the school. Nevertheless, our experiences with its implementation have been similar to those of designers of school reform programs that have a traditional focus on teaching and learning. It appears that no matter what the actual reform is, there are common factors that facilitate effective implementation. This point is made by Rand researchers (Kirby et al., 2001) in a study of schools in the New American School (NAS) initiative. They found—as we did in implementing 21C (Finn-Stevenson & Zigler, 1999)—that several factors influenced implementation, among them the principal's leadership; the school's prior experience with program implementation, teachers' perception of the need for the effort, and the support of the district. In our studies on 21C's implementation, we also found other prerequisites to successful implementation. One such prerequisite is commitment to change. Change is a characteristic of any school reform effort and indeed its goal. In the School of the 21st Century and other reform programs (see for example, Elmore, 2000), change is unlikely to occur unless there is commitment to the effort at all levels. One level of commitment is at the district level even in cases where the reform implicates only one school rather than all schools in the district (Fullan, 2001; Slavin, 2003). The implementation of 21C occurs on a school-by-school basis. However, because we seek to provide services to all children, the expectation is that all schools in the district will eventually implement the program and commitment in that regard is sought. The school district, headed by the school board and superintendent, is the primary organizational structure of the public school system, providing the schools with the external supports and infrastructure they need. Therefore the district's commitment is essential if reforms are to be implemented effectively (Ucelli, 2001).

Equally important is commitment to the program at the level of the school building, especially on the part of the principal. Principals provide not only essential leadership and support but are also actively involved in the implementation process and facilitate it. Studies on the implementation of 21C have shown that schools where principals report spending 10%–20% of their time during the first year on the program are more successful in implementation and sustainability than schools where principals report spending 5% or less of their time (Finn-Stevenson, Linkins, & Beacom, 1992). There must be a point person to assume daily coordination of 21C, but it is the principal who sets the tone for the reform, encourages change, and initiates and maintains enthusiasm and high levels of effort.

Also important is a locally driven approach to implementation. Although, as we discuss below, there are external factors that can facilitate implementation, change cannot be imposed upon a school but should be initiated from within. A locally driven approach to implementation is embedded in the design of 21C. This means that the program is the responsibility of educators in the school and is adapted to the needs of the school and community. 21C does not mandate the provision of all service components that make up the program or a uniform method for its implementation. Rather, it provides a blueprint for action and requires schools to develop and implement services on the basis of the needs of the community. The result is variations among 21C schools around the country. However, all 21C schools share a common goal: the optimal development of children through the provision of child care and support services, and they adhere to the guiding principles described earlier. Although variations in scope of effort are noted, the majority of the schools grow to provide all of the core services of the program as well as additional services implemented in response to need and requests by parents.

Implementation of a school-based program should be phased in over a period of years. This is essential in 21C, given the numerous services that make up the initiative, but it is also important for other reform efforts, as it provides a strong foundation upon which growth becomes possible (Fullan, 2001). In 21C, the decision about which services to begin with and whether or not to add additional services is made on the basis of a plan of action. The plan of action is part of the initial planning process, but it remains important in later years as well and evolves and changes over time. It is made on the basis of the following:

1. An assessment of the needs of families and an inventory of services in the community
2. An organizational audit, the purpose of which is to determine what strengths exist and what financial and other resources and capabilities the school district and individual schools have and what else is needed to facilitate implementation

Educators' ability to anticipate and address potential problems is another factor that influences implementation. When schools adopt any reform initiative, they do so because they want to bring about change, focusing on ultimately enabling students to do better academically. However, when effective implementation of an initiative takes place, other changes occur as well in the way the school operates and in the relationship between parents and educators and among educators within the school (Desimone, 2000; Fullan, 2001).

In 21C, some of the changes are logistical (for example, changes in the use of available space, change in the school's hours of operation), some are financial, and there are also numerous changes associated with staff relations. When schools are providing school-based services such as child care, there is a need to hire additional staff. Besides policy decisions regarding staff qualifications, schools have to make an effort to treat the entire staff as a whole rather than seeing the child-care staff as separate from the academic faculty. Such separation, which means, in effect, that the various child and family support services are being considered merely as an add-on, can stand in the way of the program's becoming part of the very core of the school's operation (Jehl & Kirst, 1992). The support services can indeed function autonomously within the buildings, simply using the physical space made available in the school, with no attempt made to conceptualize the programs as part of the school. However, programs are more likely to realize their potential and have positive outcomes if they are an integral aspect of the school's operations and culture.

Conceptualizing the provision of early care and family support as an integral facet of the school would lead to such activities as having representation of the entire staff early in the planning process and making provisions for regular whole-staff meetings and orientations. This helps in facilitating interactions among the staff as well as instilling the notion that the academic as well as the nonacademic support services are important functions of the school as a whole.

Fullan (1992, 2001) underscores the importance of whole-school orientation and staff development workshops, noting that these should be aspects of any school reform program. He emphasizes, however, that even when all players are enabled to participate early on in the process of change, not all of them would necessarily understand the significance of the effort or be entirely supportive of it immediately. He recommends that provisions for continued staff development be made. Huberman (1992) makes a similar point in describing several educational reform programs. He notes that initially, staff commitment to the programs was fragile, in part because not everyone understood the significance of the efforts and their relevance: "It was only when teachers had undergone a few cycles of experimentation . . . that they got on top of the [programs] in conceptual terms. This, in turn, strengthened their technical mastery and heightened their commitment" (p. 10).

Widespread Implementation and Sustainability

The ultimate goal in school reform is, of course, to move beyond the implementation of the program in a few schools and scale up the effort beyond the individual school and even the individual district. Although programs are often successful in moving from the concept stage to implementation in one or several schools, the development of programs with widespread implementation has been difficult to achieve. This point is made by Schorr and Schorr (1988), Schorr (1997), and Elias (1997), who contend that although effective programs and services exist, they often represent isolated instances of excellence and are not widely implemented. This problem is not new. At issue are not only difficulties in scaling up but also, in some cases, lack of consistency and failures in implementation when programs are replicated (Elias, Gager, & Leon, 1997; Fullan, 2001). There is also the problem of poor-quality implementation or, over time or in different places, implementing the program in such a way as to dilute the intervention. All reform programs face this problem, but early care and family support services such as those offered by 21C are especially vulnerable; financial constraints can lead to program changes or to cutting corners, thus not only diluting the effort but possibly also resulting in the provision of poor-quality programs that impede rather than enhance children's development.

The failure to implement programs to scale or having widespread implementation at a cost—namely, poor-quality implementation—is common. In part, difficulties in scaling up are attributed to the way programs are developed, with far more attention paid to the idea and initial design of a program than to its implementation beyond the model demonstration or pilot site. That is, program developers tend to put all their energies and resources into one school to demonstrate the effectiveness and feasibility of the initiative, evaluating the effort, and claiming success, only to find later that such success is limited to the demonstration site. Many of the conditions that factored into the success of the program—foremost among these being the constant attention of program developers and the coaching, nurturing, and resources they provide to educators during the implementation process—cannot be replicated. The results are efforts that do not resemble the original program and fail to produce positive outcomes (Olds, 1997).

Since making available good-quality child care was the reason for starting 21C, we help schools develop and maintain good-quality programs. A national study on 21C and another study conducted on 21C schools in Arkansas has shown that schools have succeeded in that regard, with the majority scoring high marks on the widely used quality measure known as the Early Childhood Environmental Rating Scale (Desimone, Payne, Fedoravicius, Henrich, & Finn-Stevenson, 2004; Hendrich, Ginicola, & Finn-Stevenson, 2006).

In conceptualizing the school of the 21st Century, our intent at the outset was to ensure widespread implementation. Rather than limit our work to one or several pilot sites, we took a service-oriented approach to developing 21C that entailed building our staff capacity to enable us to work with multiple schools. The need to provide schools with assistance in the implementation of reform programs is emphasized by Ucelli (2001), who notes that schools have few if any places to turn to for support in implementing reforms. Berends and colleagues (2000) also found that implementation is effective *only* in instances where schools have access to continued assistance from program designers.

Technical Assistance and Training

Our approach to providing schools with support in implementation is twofold: the provision of technical assistance and the establishment of a national network. To facilitate our support of schools, we have developed a training protocol and provide on-site and off-site technical assistance and training on the implementation of the program. The technical assistance staff, referred to as Implementation Associates, are based at Yale and have a designated number of schools in their region; they have been working with each of the schools intensely for about 2 years. As the schools become more proficient and implementation is under way, the associates assume an advisory role and provide only occasional assistance.

The implementation staff is assisted by superintendents, principals, program coordinators, and others who have successfully implemented the School of the 21st Century. We have selected from among them several peer trainers and have found that pairing them with their counterparts in schools that are beginning the implementation is an effective and efficient training approach. It not only enables educators to learn from one another but also provides an informal support system, which can be an invaluable source of strength during the initial phases of implementation. This peer approach to training is especially effective when we pair educators from similar districts or those likely to experience similar problems.

National Network

The second aspect of our effort to facilitate widespread implementation is the establishment of a School of the 21st Century National Network. 21C schools establish informal relationships and help each other in implementation. However, as the program continues to grow nationally, it has become imperative to ensure that local services maintain a high level of quality and that we are able to address the training needs of educators at established sites—hence the creation of the National Network. This entails the development of standardized professional materials and services, the codification of procedures for working with schools with varying levels of training needs and/or schools that are experiencing changes (for example, an increase in the number of immigrant children), and convening regional and national conferences.

These approaches to widespread implementation have enabled us to respond to the interest schools have in implementing the program as well to maintain enthusiasm for the effort once the program has been implemented. This national perspective is an essential ingredient of scale-up efforts that have been adopted by other initiatives, such as

Success for All (Slavin, Dolan, & Madden, 1994). However, in and of itself it is not sufficient. Also important is an understanding of the local level, described earlier, and the practical information—what difficulties to expect, how to overcome these, and what specific steps to take—about program implementation (Fullan, 1992). Such information can be obtained only when program developers work with several schools, when they immerse themselves in the local settings and contexts, and when they extend their ranks through participation in implementation (Elias, 1997, p. 261).

Our efforts to date have led to several important achievements. First, more than 1,300 schools have implemented the program and in many cases have operated 21C for two decades. Second, Connecticut and Kentucky appropriate funds for statewide implementation, where 21C's program is offered in Family Resource Centers. Third, our work with both private foundations and state agencies in Arkansas has resulted in the implementation of 21C in schools throughout the state as well as the establishment of a 21C Network and Leadership Council. The latter is designed to continue to develop 21C in the state as well as to serve as a template for the development of 21C in other states.

MAKING A DIFFERENCE

Program growth is only one measure of success. It is also essential to evaluate programs, including not only outcome studies but also process studies to determine what has been implemented and how. Several process studies have been conducted in 21C schools, the findings of which—many of them reported above—have informed our efforts and facilitated the growth of the program.

Outcome studies are more difficult to conduct and it is important to wait for programs to be fully implemented before undertaking these. There are other limitations as well, including the inability to utilize a rigorous experimental design and random assignment. This limitation is not unique to the evaluation of 21C, nor are other evaluation difficulties. Even in programs that focus on one outcome, such as increased reading scores, and are easy to implement and evaluate, difficulties in evaluation are found. In 21C and other comprehensive school- and community-based reform initiatives, there are multiple services and goals, the achievement of which depends upon interactions throughout the system. Such programs, as we noted in the previous pages, are not implemented all at once but rather over time; even when stable, they are characterized by a constant process of adaptation and change. The complex, flexible, and evolving nature of the programs render evaluations difficult, and these difficulties are exacerbated by the breadth of the range of outcomes that comprehensive programs pursue (Brown, 1995; Kubisch, Weiss, Schorr, & Connell, 1995).

These issues can be addressed and the program evaluation can be enhanced by employing, as we did, what Weiss (1995) calls "theories of change"—how and why an initiative works—as the basis for evaluation. This necessitates defining the underlying assumptions behind the cause-and-effect relationships that programs wish to achieve and developing data collection and analyses to track the unfolding of the assumptions. We are cautioned, however, that it is usually not effective to investigate a wide range of variables (Miller, 1988); in our evaluation, therefore, we had to be specific and focus only on some of the expected outcomes.

There are other difficulties, such as the need to conduct evaluations over a long period of time in order to document positive outcomes. Zigler, who has long held the view that there is no quick fix, and Brooks-Gunn (2000) note that if we expect positive outcomes after a year or two of program participation, we may be engaging in "magical thinking." Among other difficulties, schools providing 21C may also provide other reforms, such as literacy programs, sometimes putting these in place after the evaluation has begun. This presented obvious problems regarding attribution of effects and necessitated (a) careful selection of districts to participate in the study and (b) the need to include in the evaluation not only the documentation of outcomes but also the context.

Our recent work in Arkansas, where we worked with several cohorts of schools implementing 21C over a period of several years, presented an opportunity for a naturally occurring experiment at two levels. One, we studied 21C schools implementing the program at different times and found greater program impact among the more mature school, where 21C was associated with decreased rates of grade retention and behavior problems and increased rates of academic achievement, as indicated in improved reading and math scores (Ginicola, Yekelchik, & Finn-Stevenson, 2007). Two, state funds for preschool in Arkansas—under a funding program known as Arkansas for Better Care (ABC)—were given to schools. One group of schools provided child care (we refer to this as the ABC-only group). Another group of schools provided ABC-funded preschool in addition to other components of 21C (we refer to this group as 21C/ABC). In order to determine if 21C has added value, Ginicola and colleagues (2009) compared preschools that had both ABC and 21C services with those that contained only ABC services. Results indicated that at baseline, children in 21C/ABC scored the same or less well than their ABC counterparts. However, at the end of the year, the participants at the 21C/ABC preschools were significantly ahead of children in ABC-only programs on all developmental indicators.

CONCLUSION

The potential of 21C to improve academic achievement is noted in the above studies and will be used to further the growth of the program and its implementation in more schools. Our goals when we first conceptualized the program were to ensure the optimal development of all children through the provision of child care and family support and to explore the feasibility of using the schools for the provision of such services. We have shown in this chapter that this is possible. The contributions of 21C extend beyond the idea of the use of the school for the provision of child care and family support. Also critical has been our work on implementation, moving the program from idea to practice. We had to become immersed in the implementation process, and we learned from 21C schools and from the successes and failures of other school reform initiatives. Our understanding of program implementation has grown in the process of our work, underscoring the fact that there is reciprocity in the link between knowledge and practice; knowledge and theory are important aspects of school reform and, in the same way, what we learn from the process of program implementation enriches the theory and adds to our knowledge base.

NOTE

1. Some school-based programs such as Full-Service Schools and the Beacon Schools initiative simply use the school space.

REFERENCES

Barnett, W. S., Hustedt, J., Friedman, A., Boyd, J., & Ainsworth, P. (2008). *The state of preschool 2007.* Princeton, NJ: National Institute for Early Education Research.

Berends, M., Kirby, S., Naftel, S., & McKelvey, C. (2000). *Implementation and performance in new American schools: Three years into scale up.* Santa Monica, CA: Rand.

Brooks-Gunn, J. (2000). Do you believe in magic? What we can expect from early childhood intervention programs. *Social Policy Report, Society for Research in Child Development, 17,* 1–14.

Brown, P. (1995). The role of the evaluator in comprehensive community initiatives. In Connell, J., Kubisch, A., Schorr, L. & Weiss, C. (Eds.). *New approaches to evaluating community initiatives* (pp. 201–225). Washington, DC: Aspen Institute.

Consumer Product Safety Commission (1999). *Safety hazards in child care settings.* Washington, DC: U.S. Consumer Product Safety Commission.

Desimone, L. (2000). *Making comprehensive school reform effort work.* New York: Clearinghouse on Urban Education, Teachers College.

Desimone, L., Payne, B., Fedoravicius, N., Henrich, C., & Finn-Stevenson, M. (2004). Comprehensive school reform: an implementation study of preschool programs in elementary schools. *Elementary School Journal, 104,* 369–374.

Duncan, G., & Brooks Gunn J. (Eds.). (1997). *Consequences of growing up poor.* New York: Russell Sage Foundation.

Elias, M. (1997). Reinterpreting dissemination of prevention programs a widespread implementation with effectiveness and fidelity. In R. Weissberg, T. Gullotta, R. Hampton, B. Ryan, & G. Adams (Eds.), *Establishing preventive services* (pp. 283–289). Thousand Oaks, CA: Sage.

Elias, M. J., Gager, P., & Leon, S. (1997). Spreading a warm blanket of prevention over all children: Guidelines for selecting substance abuse and related prevention curricula for use in the schools. *Journal of Primary Prevention, 18,* 41–69.

Elmore, R. (2000). *Building new structures for school leadership.* Washington, DC: Albert Shanker Institute.

Fight Crime/Invest in Kids. (1997). *After school crime or after school programs: Report to the U.S. Attorney General.* Washington, DC: Fight Crime/Invest in Kids.

Finn-Stevenson, M., Linkins, K., & Beacom, E. (1992). The school of the 21st century: Creating opportunities for school-based child care. *Child and Youth Care Forum, 21*(5), 335–345.

Finn-Stevenson, M., & Zigler, E. F. (1999). *Schools of the 21st century: Linking child care and education.* Boulder, CO: Westview Press.

Fullan, M. G. (1992). *Successful school improvement. The implementation perspective and beyond.* Philadelphia: Open University.

Fullan, M. G. (2001). *The new meaning of educational change.* New York: Teachers College Press.

Garmezy, M. (1985). Stress resistant children: The search for protective factors. In J. E. Stevenson (Ed.), *Recent research in developmental psychopathology.* Oxford, UK: Pergamon.

Gilliam, W. S., & Zigler, E. F. (2001). A critical meta-analysis of all impact evaluations of state-funded preschool from 1977 to 1998: Implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly, 15,* 441–473.

Ginicola, M. M., Finn-Stevenson, M., Zigler, E., & Lapin, B. (March, 2009). *Added value of the school of the 21st century when combined with a quality statewide preschool program.* New Haven, CT: The Zigler Center in Child Development and Social Policy, Yale University.

Ginicola, M. M., Yekelchik, A., & Finn-Stevenson, M. (2007). *2007 Evaluation of the Arkansas school of the 21st century program.* New Haven, CT: The School of the 21st Century.

Henrich, C., Ginicola, M. M., & Finn-Stevenson, M. (2006). *The School of the 21st century making a difference: Findings from two research studies.* New Haven, CT: The School of the 21st Century.

Huberman, M. (1992). Critical introduction. In M. Fullan (Ed.), *Successful School Improvement* (pp. 1–20). Philadelphia: Open University Press.

Jehl, J., & Kirst, M. (1992). Getting ready to provide school-liked services: What schools must do. *The Future of Children, 2*(1), 95–106.

Kirby, S., Berends, M., & Naftel, S. (2001). *Implementation in a longitudinal sample of new American schools.* Santa Monica, CA: Rand.

Kubisch, A., Weiss, C., Schorr, L., & Connell, J. (1995). Introduction. In J. Connell, A. Kubisch, L. Schorr, & C. Weiss (Eds.), *New approaches to evaluating community initiatives* (pp. 1–22). Washington, DC: Aspen Institute.

LeMoines, S., Morgan, G., & Azer, S. (2003). A snapshot of trends in child care licensing regulations. *Child Care Bulletin, 28,* 1–5.

Miller, B., O'Connor, S., Stringnano, S., & Joshi, P. (1996). *Out of school time in three low income communities.* Wellesley, MA: Center for Research on Women.

Miller, S. H. (1988). The Child Welfare League of America's Adolescent Parents Project. In. H. B. Weiss & F. H. Jacobs (Eds.), *Evaluating family programs* (pp. 371–388). New York: Aldine de Gruyter.

National Education Goals Panel. (1991). *The National Education Goals Report.* Washington, DC: Author.

Olds, D. (1997). The prenatal/early infancy project: Fifteen years later. In G. W. Albee & T. P. Gullotta (Eds.), *Primary prevention works.* Thousand Oaks, CA: Sage.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Yazejian, N., Culkin, M. L., Zelazo, J., et al. (1999). *The children of the cost, quality and outcomes study go to school.* Chapel Hill: University of North Carolina Press.

Phillips, D. (1995, July 31). Written testimony. Committee on Education and the Workforce. Washington, DC: U.S. House of Representatives.

Putnam, R. D. (1995). Bowling alone: America's defining social capital. *Journal of Democracy, 6,* 65–78.

Rutter, M. (1979). Protective factors in children's responses to stress. In W. Kent & J. Rolf (Eds.), *Primary prevention of psychopathology* (pp. 130–139). Hanover, NH: University Press of New England.

Rutter, M. (1980). *Changing youth in a changing society.* Cambridge, MA: Harvard University Press.

Schorr, L. (1997). *Common purpose: Strengthening families and neighborhoods to rebuild America.* New York: Anchor Books.

Schorr, L. & Schorr, D. (1988). *Within our reach: Breaking the cycle of disadvantage.* New York: Doubleday.

Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development.* Washington, DC: National Academy Press.

Shonkoff, J. P., & Phillips, D. A. (2002). From neurons to neighborhoods: The science of early childhood development. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(5), 625–626.

Slavin, R. (2003, March 5). Converging reforms. Changing schools? Changing districts? How the two approaches can work together. *Education Week, 22*(25), 44–45, 64.

Slavin, R., Dolan, L., & Madden, N. (1994). *Scaling-up: Lessons learned in the dissemination of success for all.* Baltimore, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

Ucelli, M. (2001). *From school improvement to systems reforms.* New York: Rockefeller Foundation.

U.S. Census Bureau. (2000). *Families and living arrangements.* Washington, DC: U.S. Labor Department, Census Bureau.

Vandell, D. L., & Shumow, L. (1999). After-school child care programs. *Future of Children, 9*(2), 4–80.

Vinson, T., Baldry, E., & Hargreaves, J. (1996). Neighborhoods, networks and child abuse. *British Journal of Social Work, 26,* 523–543.

Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts.* Washington, DC: Aspen Institute.

Young, K. T., Marsland, K. W., & Zigler, E. (1997). Regulatory status of center-based infant and toddler child care. *American Journal of Orthopsychiatry, 67,* 535–544.

Zigler, E. F., Finn-Stevenson, M., & Hall, N. (2002). *The first three years and beyond: Brain development and social policy.* New Haven, CT: Yale University Press.

Zigler, E. F., Gilliam, W., & Jones, S. M. (2006). *A vision for universal preschool education.* New York: Cambridge University Press.

Zigler, E. F., Marsland, K., & Lord, H. (1999). *The tragedy of child care in America: The way forward.* New Haven, CT: Yale University Press.

Zigler, E. F., & Valentine, J. (1979). *Head start: The legacy of the war on poverty.* New York: Free Press.

This page intentionally left blank

14

A Thinking Laboratory: Perspectives of Prospective Education

ALEX KOZULIN

How can one fail to agree with the popular opinion that public education, at least in many of the technologically developed countries, is in a state of crisis? But does this mean that everything was okay in the "good old days"? Does this imply that the theory of instruction and learning has made no progress and that our understanding in this field has not advanced but just stagnated? In this chapter I will focus on four areas where, in my opinion, significant progress has been made during the last 20 years or so. If properly applied, the findings made in these areas can help public education move in a more promising direction.

The first area concerns the general conceptualization of the goals and framework of the educational process. This may provide a foundation for the paradigm shift from what I call "retrospective" to "prospective" education, from education oriented toward the transmission of knowledge from generation to generation to the preparation of students for a vague, uncertain, and dynamic future. A new paradigm of prospective education by necessity places greater emphasis on students' ability to orient themselves in unfamiliar knowledge fields, formulate problems, search for nonstandard solutions, and engage in self-directed learning. I will show how turning a classroom into a thinking laboratory may play a central role in this shift of the educational paradigm.

The second area that has witnessed significant progress is education in the multicultural classroom. A couple of decades ago many educators were simply unaware of the special challenges associated with the task of educating students who have different cultural backgrounds. In time, however, it became clear that the multicultural classroom is not a passing phenomenon and that what happens in it can teach us something about fundamental relationships between development and education. Cognitive and behavioral phenomena that appear as following an orderly maturational sequence, immediately reveal their culturally constructed character when observed in children belonging to different cultural groups. The growing interest in sociocultural theory has helped to conceptualize the observed cultural differences in terms of "psychological tools"—internalized and transformed symbolic tools characteristic of a given culture or subculture.

The third area is that of the evaluation of students' learning potential. Progress in this area was triggered by both dissatisfaction with the existent psychometric approaches and by inspiration received from Vygotsky's (1986) notion of a Zone of Proximal Development (ZPD). At the heart of the problem of evaluation lies a paradox clearly identified by Vygotsky: ordinary assessment methods are capable of measuring what students have already learned how to do, while what the educators need is to understand how students will learn to do things that they do not yet know how to do. One of the conceptualizations of this problem is the distinction between so-called summative and formative assessment. Summative assessment techniques aim at accurately evaluating the students' current level of knowledge and skills. Formative assessment offers a range of methodologies for helping students and their teachers adjust the teaching/learning process to the students' needs. It seems that one of the most promising developments in this area is the so-called dynamic assessment (DA) approach, which includes the learning phase within the assessment process. Using a number of examples of DA projects with different populations of learners, I will demonstrate how this methodology helps us to go beyond the simplistic dichotomy of low versus high scorers.

The last area to be explored is the emerging research on differences between such fundamental processes as thinking and learning. This problem is complicated by the fact that many authors use the notions of thinking and learning interchangeably, assuming that intelligence is a general aptitude for learning. At the same time, even in our everyday life, we distinguish between quick and efficient learners who may not be deep thinkers, and people of superior intelligence who are sometimes slow learners. Tentatively, I would suggest that learning is primarily an acquisition process, while thinking is an integration of the already acquired knowledge and skills and their application to new situations. The development of these two types of processes requires different educational methodologies. Before this is done, students with a high IQ will automatically be presumed to be good learners, while students who easily imitate models presented by teachers will be thought to possess high intellectual skills. The suggested distinction also impacts on the issue of assessment, where learning potential should apparently be distinguished from cognitive modifiability.

FROM RETROSPECTIVE TO PROSPECTIVE EDUCATION

How can one prepare children for a vague, uncertain, and dynamic future? Traditional education responded to society's important goal of transmitting an unambiguous cultural tradition from generation to generation. In this sense traditional education was retrospective. It took for granted a certain model of the world and a certain cultural tradition. Students were expected to absorb this tradition and the intellectual tools associated with it. For the most part, schools prepared students to cope with problems that reproduced the cultural patterns of the past—the problems that already had well-known solutions.

As a result, the most desirable qualities of a good student were the ability to follow the model presented by the teacher, recall and perform standard operations, and remember a large amount of factual information. Exactly the same qualities would also help a student to receive high scores on IQ tests and college admissions exams like Scholastic Aptitude Test. Do I think that these qualities are unimportant? Not at all. They are important but insufficient. In the past, the qualities beyond those of a "good student" were required for a minority of high school graduates aspiring to embark on academic careers. Scientists have never just followed the existent models, it has never been enough for them to perform standard operations, and quite often they did not have a good memory for facts but instead knew how to get information on these facts.

Now these "scientific" qualities seem to be required of almost every high school graduate. There are two main reasons for this change. The first is related to the fact that an unambiguous cultural tradition cannot be taken for granted in the era of multicultural classrooms. The second reason is an accelerated and pervasive technological innovation that turns the future into a much more uncertain and dynamic one than ever before. In a sense, educational systems are facing a paradoxical task of preparing students for future activities whose parameters are still unknown as the student is learning.

What kind of a learner will be capable of coping with such an uncertain future?

Karoly and Panis (2004) suggested the following answer in the report they prepared for the U.S. Department of Labor:

> Rapid technological change and increased international competition place the spotlight on the skills and preparation of the workforce, particularly the ability to adapt to changing technologies and shifting product demand. Shifts in the nature of business organizations and the growing importance of knowledge-based work also favor strong non-routine cognitive skills, such as abstract reasoning, problem-solving, communication, and collaboration. Within this context, education and training become a continuous process throughout the life course involving training and retraining that continues well past initial entry into the labor market.

The same themes appear in the European Commission (2006, p. 44) report on professions of the 21st century:

> Knowledge will become more ephemeral than it is today. As a result, it will be necessary to focus on teaching generic techniques as well as teaching specific knowledge. In the future, the success of an economic sector will lie in its ability to restructure itself quickly; the labor force needs to adapt to these changes as well: they will more often switch jobs or will be placed in a completely different position. Therefore, part of the workforce will have to learn to adapt to new environments. This skill is closely linked to the famous "learning how to learn."

To respond to these challenges, a new paradigm of prospective education should emphasize the following activities: learning how to orient oneself within new material; learning how to formulate and solve novel problems; and planning, organizing, and self-regulating one's own lifelong learning process.

The common denominator in various attempts to provide a theoretical basis for what I call prospective education is the emphasis on cognition. Retrospective education made a rather sharp distinction between cognition and education. Students' cognition was perceived as dependent primarily on genetics, maturation, and individual experience. Education, on the other hand, focused mostly on curricular material taking the "age-appropriate" level of students' cognition as a given. Students whose cognition was not age-appropriate were sent to special education. The continuing difficulties with students' problem solving, however, forced both researchers and educators to start looking beyond curricular material and into students' thinking skills. It became clear that one cannot wait until students' higher-order thinking skills develop "naturally"; they should be developed deliberately (Costa, 1991). This realization in itself amounted to a significant achievement in the relationship between developmental and educational theories. Children's cognitive functions ceased to be an exclusive province of developmental psychology and became an integral part of a new theory of learning/teaching that elaborated methods of actively constructing these functions. At that juncture, two main directions could be discerned. The first suggested reshaping curricular teaching and learning processes so that, instead of just providing students with content knowledge, curricular learning would also develop students' cognitive and learning functions. Probably the most consistent of these approaches is the Learning Activity model based on Vygotsky's sociocultural theory (Kozulin, Gindis, Ageyev, & Miller, 2003; Zuckerman, 2004). Already at the primary school level, this model emphasizes conceptual learning and the development of metacognitive skills. The second direction focuses on developing students' thinking skills in a pure form and only then "bridging" them to curricular material. A number of cognitive skills programs were developed (for a review, see Harpaz, 2007); probably the most comprehensive of these programs is the Instrumental Enrichment Program of Feuerstein and colleagues (1980).

In an attempt to combine the strong points of the Learning Activity and Instrumental Enrichment approaches, Kinard and Kozulin (2008) designed their Rigorous Mathematical Thinking (RMT) program. It is based on the assumption that much of students' failure in mathematics stems not from a lack of mathematical knowledge but from the underdevelopment of their general cognitive skills on the one hand and insufficient emphasis on mathematical conceptual reasoning on the other.

The RMT program has three major aspects:

1. Developing students' cognitive functions and problem-solving skills through appropriation and internalization of first general and then mathematically specific symbolic tools. Codes, tables, and diagrams are examples of general symbolic tools, while number

lines, formulas, and coordinate systems are mathematically specific tools. This is achieved by the inclusion of special cognitive tasks in the mathematics curriculum and by emphasizing the instrumental role of such mathematical tools as graphs, formulas, and so on.

2. Transforming teachers from providers of information into mediators of conceptual reasoning. This is achieved by organizing training and supervision sessions during which teachers are encouraged to think how a given classroom activity would benefit students' general cognitive development and mathematical conceptual reasoning.
3. Replacing mechanical rehearsal of standard operations with conceptually oriented problem solving. This is achieved by subordinating a study of operations to the acquisition of mathematical concepts.

The implementation of RMT starts with teachers being encouraged to select a mathematical concept or a "big idea." Then they identify those cognitive functions that are required for successful concept formation. They probe the state of these functions in their students and select cognitive tasks and symbolic tools for the development of those functions that are weak or deficient. Using the emerging cognitive functions, teachers create situations calling for concept formation. Mathematical concepts appear in the RMT program as generative: by grasping the main idea of the concept, students then create all possible empirical realizations and check the limits of the concept's applicability. Finally the newly acquired concept is connected with previously studied mathematical concepts, so that the students' knowledge is systematic rather than episodic.

The application of the program in a number of inner-city schools in the United States has demonstrated that students who participated in RMT activities made greater gains than their peers who received a standard math curriculum. The RMT advantage was observed both in cognitive tasks and in standards-based math performance exams and remained 6 months after the end of intervention (Kinard & Kozulin, 2008, chapter 6). Probably more important, however, is that the program helped even those students who previously "hated" mathematics and were considered hopeless underachievers.

Though cognitive education as a response to the challenge of prospective education looks encouraging, it is not without its problems, especially at the level of classroom implementation. For each one of the directions in cognitive education identified by Harpaz (2007), there are typical implementation pitfalls. The cognitive skills approach becomes *tamed,* losing its transformative power and becoming just another school subject that will be forgotten the minute students pass the cognitive skills exam. The curriculum-based cognitive programs lose their constructive orientation and become *lecturing* about concept formation rather than actually developing concept formation skills. These problems notwithstanding, it is difficult to imagine what kind of approach, other than cognitive education, might prepare students for an uncertain and dynamic future.

LEARNING IN A MULTICULTURAL CLASSROOM

About 15 years ago I gave a seminar on cognitive methods of teaching in a multicultural classroom as part of an international summer school. The participating teachers came from different countries, but it so happened that a considerable number were from Italy. At the end of the session, the Italian participants thanked me for the presentation but commented that, for them, it was of little practical value because there are no immigrant students in Italian schools and they do not foresee the situation to change in the future. Ten years later, I was invited to lecture on the methods of teaching in a multicultural classroom to a packed audience in a medium-sized Italian city. The participants told me that they were at a complete loss as to what to do with immigrant students from at least 10 different countries who did not speak Italian and had had very different previous educational experiences. The situation in bigger cities was even more challenging in terms of cultural variety and the number of non-Italians.

What happened during these 10 years was not only an influx of immigrant students into European and American classrooms but also a growing awareness that the multicultural classroom is not a passing phenomenon but will be a permanent feature of education in the future (Trumbull & Pacheco, 2005). One of the major challenges revealed by multicultural classrooms is the cross-cultural difference in students' problem-solving and cognitive skills. Retrospective education made a sharp distinction between children's cognitive development, perceived as a natural process, and classroom learning anchored to curricular content and the transmission of disciplinary skills. Cognitive development within this model was perceived as universal: all people were believed to have the same cognitive functions and problem-solving skills, while the quantitative differences were explained by either individual or ethnic group variation. The educational process was supposed to be based on already matured cognitive functions.

The sociocultural approach offers an alternative model of the relationships between development and education and, as a consequence, also an alternative explanation of cross-cultural differences in cognition. Human cognition is perceived as shaped by the symbolic tools that are available and the sociocultural activities typical in a given culture. Because cultures differ both in symbolic tools and in sociocultural activities, the resultant cognitive functions will also be different. It is important to realize that we are not talking about the content of culture but about the most basic cognitive mechanisms fostered in different ways by different cultures.

Let us take spatial memory as an example. At first glance, it seems that to remember the positions of objects is such a basic skill that it should be determined exclusively by the natural abilities of each individual. The research, however, has demonstrated that even this ability is shaped by symbolic tools (Kozulin, 2008a). One of the simplest ways to check positional memory is to show a person a 5-by-5 grid and then point to specific positions, as shown in Figure 14.1. After the first demonstration, the person is asked to mark the pointed positions on his or her answer sheet, which has the same 5-by-5 grid. After that, the procedure is repeated either for a set number of times or until the person has marked all the positions correctly. If positional memory depended exclusively on individual abilities, one would not expect any significant cross-cultural differences. And yet it was discovered that young adults brought up in Ethiopia, whose educational experience was limited to a few years in a rural school, experienced greater difficulties with this task than college-educated Europeans. From the responses given by the participants and the analysis of their mistakes, it became clear that Ethiopians ignored the grid almost completely as a possible tool for remembering positions. For Europeans, the grid provided a clear reference system of rows and columns, so that each position had its unique "address," for example, 1st column, 1st row; 2nd column, 5th row; and so on. For Ethiopians, the pointed positions remained detached from the grid; in a sense, the latter became a distracter rather than a tool.

In the above case, the ubiquitous presence of the grid structure in the European culture, from the grid of city streets to the graph paper of the math exercise book, supplied European students with the symbolic

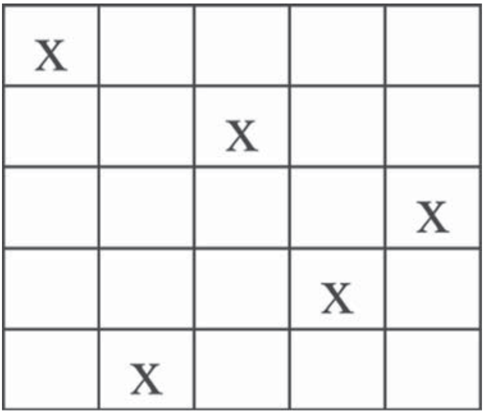


Figure 14.1 Positional memory task.

tools that shaped their spatial memorization function. If, for example, instead of examining positional memory we were checking memory for verbal puzzles presented in an oral modality, there is a good chance that Ethiopian participants would have demonstrated superior results. This is because learning strategies for the memorization of verbal puzzles is a popular sociocultural activity in Ethiopian villages but not in European cities.

The fact that certain symbolic tools are more familiar and better internalized in certain cultures does not imply that cultural novices cannot master them and in this way reshape some of their cognitive functions. In the above study of positional memory, it was shown that mediation of grid strategies to the Ethiopian students significantly improved their performance during the test. Their participation in a cognitive training program resulted in even better performance with positional memory tasks at the end of the program. It can thus be concluded that some of the cognitive functions that at first glance seem to be natural and universal are actually shaped by culturally specific symbolic tools and the person's involvement in certain sociocultural practices.

Cultural differences should not necessarily be as dramatic as are those between rural Ethiopia and urban western Europe in order to have an impact on cognitive functions. Sometimes even relatively minor variations in educational and sociocultural activities lead to the different shaping of cognitive functions. Consider the following studies, which focused on the person's ability to create instructions. The initial study was conducted by Marti and Majordomo (2001) with the aim of investigating how children of different ages use text and graphics to create instructions for their peers, who had to make a call from a particular cellular phone. The results of this study showed that notations of younger children were not very accurate and constituted continuous text without any graphic elements or other symbolic notations. Older children and adults presented a more accurate sequence of actions, using a combination of schematic drawings and text; they also often shaped their instructions as a list with headings 1, 2, and 3. Only about 50% of 9- and 10-year-old children used both text and graphics in their notations, while about 80% of adults used such a combination.

Marti and Majordomo interpreted their findings as demonstrating an orderly maturational progression of the cognitive function of representation. From my point of view, however, cognitive changes that the authors associated with maturation might reflect the progressive acquisition of specific symbolic tools that became internalized and then deployed in the task of creating instructions. Adult participants of the Marti and Majordomo study were undergraduate university students. During their high school and university studies, they certainly were trained to use a variety of notational systems including lists, tables, schemas, diagrams, and so on. To test the idea of cultural influence, I replicated Marti and Majordomo's study with women educated in ultrareligious Jewish schools and colleges (Kozulin, 2008b). In these schools, such subjects as mathematics are taught only at a basic level, while science is usually not taught at all, the bulk of the time being devoted to religious subjects. The results of my study showed that only 31.6% of the participants used a combination of text and graphics—a rate lower than that of primary school students in Marti and Majordomo's research. At the same time, their use of the list format in textual instructions was similar to that of the adults in Marti and Majordomo's sample.

From the maturational perspective, these results are paradoxical. When judged by the criterion of combination of text and graphics, the women in my study are supposed to be at the developmental level of a 9- or 10-year-old; while when judged by their use of the list format, they are at the adult developmental level. The sociocultural perspective allows for an alternative interpretation of these results. Instead of presuming an orderly maturational progression from one form of representation to the next, we should focus on those forms of sociocultural activity the students are involved in and those symbolic tools that they had a chance to appropriate. Jewish religious education for females provides little opportunity for appropriating such symbolic tools as diagrams or charts, which constitute elements of the science curriculum. At the same time, religious texts and commentary use all sorts of list notations. It would appear that the form of notation used by the participants was directly related to those symbolic tools that they had acquired and internalized in the course of their education.

These findings have direct implications for teaching/learning in a multicultural classroom. The students' cognitive functions reflect not only their maturational level but also, probably to an even greater extent, their experience with specific symbolic tools and sociocultural activities. The differential development of cognitive functions can be conceptualized through the notion of psychological tools suggested by Vygotsky (1930/1979). The formation of psychological tools starts with the acquisition of certain symbolic systems available in a given community. Among these systems are signs, symbols, oral and written languages, graphic organizers, pictorial images, and so on. Each culture and subculture has its own system of symbolic tools. Already at this level there can be considerable diversity. As shown above, certain subcultures may lack tools that are readily available in cultures based on literacy and formal schooling, and vice versa. However, it is not enough for the symbolic tools to be available; they should actually be appropriated by students as tools, that is, as instruments of learning and problem solving. For many students, symbolic systems are perceived as a part of content information rather than as tools. As a result, problem solving based on the use of such a simple tool as a table, for example, remains deficient.

In the study of Schur and Kozulin (2008), 13- to 14-year-old students from middle-class cultural majority families were given a task that required comparing the properties of different stars listed in the table and identifying two correlated properties. No knowledge of astronomy was required for solving this problem; it was enough to make systematic comparisons by using a table as a tool. In spite of its apparent simplicity, more than 40% of the students failed to complete this task correctly. This seems to be a case where a table remained an external artifact that was quite familiar to students but which failed to become a symbolic instrument of problem solving.

The appropriation of symbolic systems as external tools of learning and problem solving is only the first stage. Then these systems should be internalized to become inner psychological tools. The process of internalization and transformation of external symbolic systems into inner cognitive functions is not trivial even for educated adults (Kozulin, 2005a). Teachers who participated in continuing education training were given

24 numbers and asked to classify them as odd/even and by the number of digits (one-, two-, and three-digit numbers) and present the result of classification as a table or a chart. Only 48% of them spontaneously selected the optimal form (two columns and three rows). This result points to the failure of internalization. Teachers who had no difficulty using tables as external organizers showed a deficiency in spontaneously thinking about numerical data in a tabular form.

The problems with internalization of more specialized systems of symbolic tools—such as those used in physics, chemistry, computer programming, and so on—underscore the fact that every modern school is "multicultural" in the sense that all students irrespective of their native language are expected to appropriate and internalize symbols and languages of different areas of knowledge. The notion of psychological tools, which first appeared to be uniquely advantageous for explaining cognitive differences in the multicultural classroom, seems to have a much more general applicability as a guiding principle of appropriation and internalization of any "culture" and any "literacy."

HOW SHOULD STUDENTS' POTENTIAL BE EVALUATED?

> Imagine two seven-year-old children who showed exactly the same results in IQ tests. Now let us start helping them to solve more difficult tasks by asking questions, providing cues, or showing models. With this help, the first of them advanced to the level corresponding to a mental age of nine, while the second reached just seven years and six months.

This is how Vygotsky (1935, p. 13) illustrated the difference between already formed abilities and the emerging functions belonging to the Zone of Proximal Development (ZPD).

The challenge of evaluating students' potential is that ordinary assessment methods are capable of measuring only current performance, while what educators need is to know how students will learn new things. Though there have always been researchers who attempted to evaluate learning potential rather than current performance, their work was almost completely overshadowed by the dominant psychometric paradigm (for a historical review, see Kozulin, 2005b).

Some recent developments in the field of educational research indicate, however, that a certain change in this balance of power is taking place. The first of these developments is the growing popularity of so-called formative assessment (Black & Wiliam, 2003). Typical school exams are oriented toward the present; they take stock of what students already know and grade them accordingly. Formative assessment is future-oriented because its aims at helping students to improve their learning by providing them with elaborate feedback on their strong and weak points, which can also be used by their teachers for more efficient educational intervention. Though there are many types of formative assessments, practically all of them include active interaction between students and teachers. Learning interaction thus becomes an integral part of the assessment process.

The second development is associated with the Response to Intervention (RTI) approach (Fuchs & Fuchs, 2006; Grigorenko, 2009). The notion of RTI provides a conceptual basis for early intervention with children who experience reading difficulties and for identifying students with learning disability. At the core of RTI lies the belief that evaluation of the children's response to intensive intervention provides more relevant information about their learning needs than any static achievement tests. As in the case of Formative Assessment, the emphasis is not on what students can or cannot do but on how they respond under conditions of active learning mediated by teachers.

In the field of cognitive assessment, there is also a growing recognition that standard psychometric testing should be, if not replaced, then at least supplemented by assessment of students' learning potential (see Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). The central principle of the dynamic assessment (DA) approach reflects a rather obvious but rarely fully comprehended difference between performance ability and learning ability. Some aspects of this difference are captured in Vygotsky's notion of the ZPD (Chaiklin, 2003). Vygotsky argued that standard psychometric assessments are capable of detecting only those abilities that are already fully formed and can be displayed by children under conditions of solitary test taking. In his opinion, one can detect as-yet-unformed but emerging abilities under conditions of joint problem solving involving a student and a more competent mentor. Since Vygotsky's time, DA has acquired its own methodology by inserting an active learning interaction into the procedure of cognitive assessment (Feuerstein, Rand, & Hoffman, 1979; Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). DA research confirmed that learning potential is indeed different from performance: some students who score lower on static psychometric tests show significant learning gain, while some of their peers who score higher on the static pretest demonstrate a weaker learning gain.

One of the possible applications of the DA approach is to provide more reliable guidance for cognitive intervention. A study of primary school students' performance with Raven matrix problems showed that those who performed better at the DA assessment conducted at the beginning of the school year also showed better results at the end of the 9-month cognitive intervention (Kozulin, 2005c). The regression analysis confirmed that DA scores are a much better predictor of postintervention performance than static pretest scores.

This finding is not confined to young learners. In another study, 18-year-old adults of Ethiopian origin who showed poor results on standard psychometric tests were given an opportunity to participate in an alternative DA and subsequent intensive cognitive training (Kozulin, 2006). The predictive power of DA was investigated by comparing postintervention scores of subjects matched on their pretest scores on Raven Standard Progressive Matrices who had contrasting DA scores. In other words, both subgroups had the same average static pretest score, but one had a significantly higher learning potential than the other. When the postintervention scores were compared, the subgroup with higher learning potential demonstrated significantly better results. The effect size was large (d = 1.27). Learning potential as established with the help of DA is therefore a more sensitive indicator of the impact of cognitive training than a standard psychometric performance score. The practical implication is that the type and amount of intervention should be attuned to the students' learning potential rather than their initial performance level.

The potential usefulness of the DA approach extends beyond cognitive performance into curricular areas. In a study of Kozulin and Garb (2002), academically at-risk college students were given DA of English as a foreign language. The procedure included a standard reading pretest, a learning phase focusing on the analysis of the students' pretest mistakes and development of more efficient reading comprehension strategies, and a parallel posttest. The results of this DA were used to fine-tune the instruction. Figure 14.2 shows the pre- and posttest results of four students. Students T. and H. both had 29% correct answers at the pretest, but while T. made a significant gain and reached the level of 59% of correct answers at the posttest, H. advanced to only 38%. This is also true for students A. and L., who had a higher initial performance level of 62%. A. progressed to 82% of correct answers, while L. stayed at the pretest level. The practical recommendations based on these results were as follows: student A. was recommended for transfer to a more advanced group, L. was given additional exercises for the use of reading comprehension models, T. was encouraged to tackle more advanced English material, and H. was given a series of nonverbal and Hebrew (native language) cognitive exercises to develop her general problem-solving skills.

Though DA of language and reading both in L1 (first language) and L2 (second language) is a relatively recent phenomenon, it seems to be gaining in popularity. It was shown to be beneficial for minority English language learners (Pena et al., 2006), students of foreign languages (Poehner, 2008; Sternberg & Grigorenko, 2002), and children with serious language problems associated with Down's syndrome (Alony & Kozulin, 2007). One may look upon DA in curricular areas as a more structured form of formative assessment.

What then prevents formative assessment or DA from assuming a more prominent position in educational research and practice? One reason is that assessments, especially in a high-stakes situation, are externally rather than internally oriented; they are used for selection, placement, control, and accountability rather than the promotion of learning. The interests of policy makers and test publishers are often at variance with the interests of teachers (Schoenfeld, 2007). Another reason, however, is that there still seems to be some uncertainty about the ultimate goal of

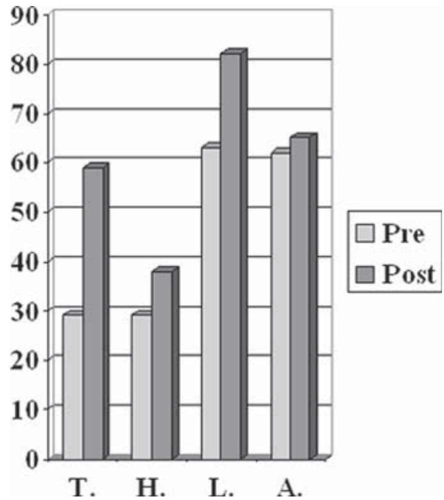


Figure 14.2 Pre- and posttest reading comprehension results of four students.

DA: does it aim at identifying students' learning potential or their cognitive modifiability? This, in its turn, is related to a lack of clear demarcation between thinking and learning. A possible approach to this problem is discussed in the next section.

THINKING AND LEARNING

The relationship between thinking and learning has been complicated by a multitude of overlapping concepts, such as "intelligence," "cognition," "experience," "problem solving," and so on. Moreover, there is a clear tendency by some authors to use the terms *thinking* and *learning* interchangeably and even to define intelligence as a "general aptitude for learning" (Slavin, 1994, p. 134). Yet even in our everyday language we tend to distinguish these two notions. One may say about a certain person: "She is a quick learner but not a great thinker," or "He is such an intelligent person but a poor learner." Grigorenko and Sternberg (1998, p. 92) seemed to express a similar view when they observed that "it is quite common to find children who show good cognitive test results but have a slow rate of learning and vice versa."

In their study of learning and problem solving in primary school children, Brown and Ferrara (1985) demonstrated that children with a higher IQ are not necessarily better learners. Students were classified as "slow" or "fast" learners on the basis of the median number of prompts required for solving the letter-sequence problem. If thinking ability as reflected in the IQ scores were indistinguishable from learning ability, then all "high IQ" students would have been "fast learners" and all "average IQ" students, "slow learners." The results demonstrated, however, that "a good third of the children had learning speed not predictable from their IQ scores" (Brown & Ferrara, 1985, p. 288). In other words, though the majority of "high IQ" students indeed turned out to be fast learners, some demonstrated slow learning. At the same time, some of the fast learners showed lower IQs. Moreover, when confronted with tasks that required integration of a number of previously learned strategies, some of the fast learners needed more prompts than some of the slow learners. Thus, learning to integrate strategies seems to be different from both direct learning from the model and from problem solving assessed by IQ tests.

Somewhat similar results were obtained more recently in a study of the thinking and learning skills of a relatively large group (N = 88) of primary school immigrant students (Kozulin, 2007). Students' learning ability was operationalized as a DA score, with tasks constituting a variation of those Raven matrices that required analogical reasoning. The DA procedure included mediating strategies for the efficient solution of a sample analogy task and then presenting students with a series of increasingly difficult tasks of the same type for independent solution. Students' thinking ability was evaluated with the help of the Raven Colored Matrices test given a week after the DA. The range of cognitive function required for solving Raven Colored Matrices is wider than just analogical reasoning. If learning and thinking abilities are the same, all students with an above-median score in DA should also show the above-median results in Raven test. The chi square test was performed to test the hypothesis that no more than 10% of the students would belong to the mixed groups—that is,"good learners but poor thinkers" and "poor learners but good thinkers." The results (chi square = 22.12, $df = 3$, $p < .001$) indicated that the hypothesis should be rejected. Actually, 27.3% of students belonged to these mixed groups. In other words, for more than a quarter of the students, efficient use of learned strategies did not guarantee efficient independent problem solving in the future, and vice versa.

How can one conceptualize this disparity? Learning might be conceived of predominantly as acquisition and thinking predominantly as application. In the acquisition process, the major emphasis is on a learner's ability to use examples, models, and cues for the acquisition of rules, concepts, and skills. Quick learners need fewer examples and are more skillful in using models and cues than slow learners. On the other hand, problem solving mostly depends on the ability to select, integrate, and apply one's existent knowledge and skills to a given problem. An intelligent person may have the same basic skills as other people but be more efficient in selecting and deploying those skills that are most relevant for solving a given task.

The implication of thinking/learning differences for the assessment of students' potential is two-fold. First, before beginning an assessment, one must decide on a goal. If the goal is to determine the students' degree of efficiency in using examples, models, and cues for solving a given set of problems, one should create a typical situation for assessing learning potential. If, on the other hand, the aim is to discover what condition or intervention would result in the students' changing their level of reasoning, then situations designed to assess cognitive modifiability should be constructed. The assessment technique should be attuned specifically to the intended goal in each case. Currently, procedures intended to assess learning potential rely heavily on cues. The use of models, examples, or partial solutions is infrequent. All these forms of mediation should be explored because they would provide us with additional information regarding students' mastery of different learning prompts. The procedure for assessing cognitive modifiability is of a different type. For example, Karpov (2008) showed that an interactive assessment procedure helps to identify young children who are ready to make a transition from the sensorimotor level of problem solving to the level based on visual imaging and eventually to the verbal symbolic level. Feuerstein et al. (1979) also used DA to ascertain students' transition from the level of trial and error to the level of hypothesis testing. In both cases the procedures focus on qualitative change in the level of problem solving rather the efficiency of learning strategies for completing essentially the same types of tasks. Many of the Program for International Student Assessment (PISA) tasks aim at assessing students' thinking rather than learning ability; the only problem is that the PISA assessment procedure is not dynamic.

Educational practice requires both types of DA. In its ideal form, an assessment of learning potential that includes a variety of prompts would provide information on the optimal form of mediation for the enhancement of students' strategies in a given type of problem. For example, instead of just providing a summative assessment of students' skills in solving arithmetic problems, a mathematical learning potential exam would generate information on both the efficiency of students' learning and their mastery of specific prompts or mediational interactions. On the other hand, a mathematical modifiability exam would evaluate students' readiness to benefit from intervention in making a transition from arithmetic to algebraic reasoning.

SUMMARY

This chapter opened with a question regarding those areas in which the theory and practice of learning and instruction can claim significant progress in the last couple of decades. Three such areas were identified: cognitive instruction, education in the multicultural classroom, and dynamic assessment. All these developments have a common denominator in the recognition that without serious attention to students' cognitive functioning, education would not be able to respond to the challenges of rapidly changing technological environment. Thus cognitive instruction must be made an integral element of the educational process. This can be achieved by adding training in thinking skills to the curriculum or/and reshaping curricular teaching in such a way as to give it clearly defined cognitive goals. My personal preference is a combination of both approaches. Some students may first need to establish general cognitive skills with the help of specially designed tasks, while others would benefit more from cognitive enhancement through a content-based curriculum. In addition, curricular teaching itself should become aimed at the development of a system of conceptual understanding rather than a collection of separate skills and operations.

Though cognitive enhancement is important for all students, it becomes crucial under conditions of a multicultural classroom. The concept of psychological tools helps to overcome both the radical relativism that rejects the notion of cognitive functions essential for formal learning and the simplistic interpretation of cultural difference as "deficiency." Because students come to the classroom with different psychological tools rooted in their home culture, none of the prerequisite cognitive functions can be taken for granted. These functions should be systematically developed through an educational process that includes higher forms of domain-specific conceptual reasoning. The experience of multicultural classrooms has implications beyond ethnic and linguistic differences. It shows that all students should be provided with symbolic tools and the techniques of their internalization, thus preparing them for being active in the various "cultures" of science, mathematics, and the humanities.

Effective education is impossible without updated forms of assessment. For a long time students' assessment was restricted to summative tests that presented a more or less accurate picture of students' current achievements but said little about their potential. First, this dominance of static, summative tests was challenged in the area of cognitive functions, where a clear distinction was shown between current cognitive performance under the condition of solitary problem solving and learning potential demonstrated under the condition of joint activity with a mentor. More recently the principles of dynamic assessment was extended to curricular areas such as reading and the learning of a second language. Today curricular formative assessment has become a recognized form of educational evaluation aimed at promoting students' progress rather than just recording their scores. It is important, however, not to lose the connection between the cognitive and curricular aspects of the evaluation procedure. An elaboration of the thinking/learning difference offered in the above discussion may help to identify a proper place for various dynamic assessment techniques

REFERENCES

Alony, S., & Kozulin, A. (2007). Dynamic assessment of receptive language in children with Down syndrome. *Advances in Speech-Language Pathology, 9,* 323–331.

Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal, 29,* 623–637

Brown, A., & Ferrara, R. (1985). Diagnosing zones of proximal development. In J. Wertsch (Ed.), *Culture, communication, and cognition: Vygotsky's perspectives* (pp. 273–305). New York: Cambridge University Press.

Chaiklin, S. (2003). The zone of proximal development in Vygotsky's theory of learning and school instruction. In A. Kozulin, B. Gindis, V. S. Ageyev, & S. M. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 39–64). Cambridge: Cambridge University Press.

Costa, A. (1999). *Developing minds: A resource book for teaching thinking.* Alexandria, VA: Association for Supervision and Curriculum Development.

European Commission. (2006). *Changing professions in 2015 and beyond.* Brussels: EC Directorate-General for Research.

Feuerstein, R., Rand, Y., & Hoffman, M. (1979). *Dynamic assessment of retarded performer.* Baltimore, MD: University Park Press.

Feuerstein, R., Rand, Y., Hoffman, M., & Miller, R. (1980). *Instrumental Enrichment.* Baltimore, MD: University Park Press.

Fuchs, D., & Fuchs, L. (2006). Introduction to response to intervention: What, why, and how valid is it?" *Reading Research Quarterly, 44*(1), 93–99.

Grigorenko, E. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities, 42*, 111–132.

Grigorenko, E., & Sternberg, R. (1998). Dynamic testing. *Psychological Bulletin, 124,* 75–111.

Harpaz, Y. (2007). Approaches to teaching thinking: Toward a conceptual mapping of the field. *Teachers College Record, 109*(8), 1845–1874.

Haywood, C., & Lidz, C. (2007). *Dynamic assessment in practice.* New York: Cambridge University Press.

Karoly, L., & Panis, C. (2004). *The 21st century at work.* Santa Monica, CA: Rand.

Karpov, Y. (2008). Do all dynamic assessment techniques assess learning potential? *Journal of Cognitive Education and Psychology, 7*(3), 411–418.

Kinard, J., & Kozulin, A. (2008). *Rigorous mathematical thinking: Conceptual formation in the mathematics classroom.* New York: Cambridge University Press.

Kozulin, A. (2005a). *Who needs metacognition more: Teachers or students?* Annual Meeting of American Educational Research Association, Montreal.

Kozulin, A. (2005b). Learning potential assessment: Where is the paradigm shift? In D. Pillimer & S. White (Eds.), *Developmental psychology and the social changes in our time.* New York: Cambridge University Press.

Kozulin, A. (2005c). Learning potential score as a predictor of sensitivity to cognitive intervention. *Educational and Child Psychology, 22,* 29–39.

Kozulin, A. (2006). *Learning potential assessment as a tool for professional advancement of cultural minority young adults.* Northumbria EARLI Assessment Conference, University of Northumbria, UK.

Kozulin, A. (2007). *Thinking and learning: One dynamic assessment procedure or two?* Annual conference on Dynamic Cognitive Intervention, School of Health Professions, Tel Aviv University.

Kozulin, A. (2008a). Evidence of culture-dependency and modifiability of spatial memory of young adults. *Journal of Cognitive Education and Psychology, 7*(1), 70–80.

Kozulin, A. (2008b). Sociocultural paradigm. In J. Clegg (Ed.). *The observation of human systems* (pp. 9–28). New Brunswick, NJ: Transaction Books.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International, 23,* 112–127.

Kozulin, A., Gindis, B., Ageyev, V., & Miller, S. (Eds.), (2003). *Vygotsky's educational theory in cultural context.* New York: Cambridge University Press.

Marti, E., & Majordomo, R. (2001). *How to use a cell phone? Children's and adults' written instructions.* Paper presented at the conference of the European Association for Research in Learning and Instruction, Padua, Italy.

Pena, E., Gilam, R., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., et al. (2006). A dynamic assessment of school-age children's narrative ability. *Journal of Speech, Language and Hearing Research, 49,* 1037–1957.

Poehner, M. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development.* New York: Springer Publishing.

Schoenfeld, A. (Ed.). (2007). *Assessing mathematical proficiency.* New York: Cambridge University Press.

Schur, Y., & Kozulin, A. (2008). Cognitive aspects of science problem solving: Two mediated learning experience based programs. *Journal of Cognitive Education and Psychology, 7*(2), 266–287.

Slavin, R. (1994). *Educational psychology.* Boston: Allyn & Bacon.

Sternberg, R., & Grigorenko, E. (2002). *Dynamic testing.* New York: Cambridge University Press.

Trumbull, E., & Pacheco, M. (2005). *The teacher's guide to diversity: Human development, culture and cognition.* Providence, RI: The Education Alliance, Brown University.

Vygotsky, L. (1930/1979). Instrumental method. In J.Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 134–143). Armonk, NY: Sharpe.

Vygotsky, L. (1935). Problema obuchenija I umstvenntogo razvitija v shkolnom vozraste. In L. Vygotsky (Ed.), *Umstvennoe razvitie detej v processe obuchenija.* Moscow: Uchpedgiz.

Vygotsky, L. (1986). *Thought and language* (rev. ed.). Cambridge, MA: MIT Press.

Zuckerman, G. (2004). Development of reflection through learning activity. *European Journal of Psychology of Education, 19,* 9–18.

This page intentionally left blank

15

Academic Intelligence Is Not Enough! WICS: An Expanded Model for Effective Practice in School and Later Life

ROBERT J. STERNBERG

What are the qualities students need to develop in order to become active, reflective, and involved citizens and professionals who achieve success in their life endeavors? How do these qualities go beyond those that we typically foster and evaluate among students in liberal arts courses in colleges and universities? If there is a discrepancy, is it possible that we in the academy are, at some level, mispreparing students for the world in which they will find themselves? And if so, are there elements we can add to a liberal arts education that will more fully address the qualities our graduates will need for successful engagement in the world? These are the central questions I seek to address in this chapter.

THE WICS MODEL

I propose the WICS model as a possible common basis for the development of skills and attitudes in college (Sternberg, 2003, 2005, 2008b; Sternberg & Grigorenko, 2004). WICS is an acronym standing for wisdom, intelligence, and creativity, synthesized. Wisdom, intelligence, and creativity, I will argue, are sine qua nons for the citizens and professionals of the future and really for anyone who wishes to achieve meaningful success in life.

It is important to state at the outset that all of these qualities are modifiable and dynamic. One is not born with a fixed level of wisdom, intelligence, or creativity but rather develops these attributes over time. They are forms of developing expertise (Sternberg, 1998a). All of us, of course, are born with some genetic predispositions. But during the course of a lifetime, these predispositions are modified by our experience such that they are developed at different rates and with different levels of success as a function of the interaction between genes and environment (Sternberg & Grigorenko, 1999).

In the remainder of this essay, I discuss each of these attributes, although for didactic purposes I do not discuss them in the order they are given above. I start with intelligence, which is a basis for creativity and wisdom and so should be discussed first. Within this discussion, I deal first with the analytic/academic aspect of intelligence and then the practical one. Next I discuss creativity. Finally, I discuss wisdom, which builds on but goes beyond intelligence and creativity. I then describe methods for developing and measuring the attributes. Finally, I draw some general conclusions.

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 (Sternberg & Detterman, 1986). Spearman (1904) first proposed that intelligence comprises a single general ability (*g*) as well as more specific abilities. This view has been extended by Carroll (1993), who is one of a number of theorists who have proposed hierarchical models, with general ability at the top and successively more specific abilities at lower levels. Such models might, for example, distinguish among verbal comprehension ability, mathematical ability, spatial ability, and so forth.

Howard Gardner (1983, 1993b, 1999) does not view intelligence as a single construct. However, instead of speaking of multiple abilities that together constitute intelligence, like some other theorists, Gardner proposed a theory of multiple intelligences in which eight distinct intelligences function somewhat independently of one another: linguistic, logical–mathematical, spatial, musical, bodily–kinesthetic, interpersonal, intrapersonal, and naturalist. Gardner (1999) has also speculated on the possible existence of existential and even spiritual intelligences. Each intelligence is a separate system of functioning. Nevertheless, these systems can interact to produce intelligent performance. For example, novelists rely heavily on linguistic intelligence but might use logical–mathematical intelligence in plotting story lines or checking for logical inconsistencies. Measuring intelligences separately might produce a profile of skills that is broader than would be obtained from, say, measuring verbal and mathematical abilities alone. This profile could then be used to facilitate educational and career decisions. Teaching to these various intelligences would require a teacher to integrate a broad range of teaching methods, such as teaching about the history of the United States through words (linguistic), maps (spatial), and songs (musical).

In order to identify particular intelligences, Gardner used converging operations, gathering evidence from multiple sources and types of data. The base of evidence includes (but is not limited to) the distinctive effects of localized brain damage on specific kinds of intelligences, distinctive patterns of development in each kind of intelligence across the life span, evidence from exceptional individuals (from both ends of the spectrum), and evolutionary history.

Gardner's view of the mind is *modular.* Modularity theorists believe that different abilities 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 relevant portions, but hard evidence for the existence of separate intelligences has yet to be produced.

Theory of Successful Intelligence: What Is Successful Intelligence?

This essay draws upon my theory of successful intelligence. According to this theory (Sternberg, 1997, 1999), successful intelligence is the ability to achieve one's goals in life, given one's sociocultural context, by capitalizing on strengths and correcting or compensating for weaknesses in order to adapt to, shape, and select environments through a combination of analytical, creative, and practical abilities.

Consider each of the three kinds of abilities in turn. Analytical ability involves analyzing, evaluating, judging, inferring, critiquing, and comparing and contrasting. Creative ability involves creating, designing, inventing, imagining, supposing, and exploring. Practical ability involves applying, using, implementing, contextualizing, and putting into practice.

The three sets of cognitive skills are somewhat different and are often found in different people in different degrees. Consider three students (who are genuine but whose names have been changed) who motivated the theory.

Alice was the teacher's dream. Her analytical skills were superb. She scored high on tests, performed well in class, and in general did everything a teacher would expect a bright student to do. As a result, Alice was initially considered in her university studies to be at or near the top of the class. Her stellar test scores were accepted as a valid indicator of her ability to do outstanding work throughout her academic career. Yet by the time she had finished her studies, she was performing at a very modest level. About 80% of her classmates were doing better than she was. What went wrong? The answer, quite simply, is that whereas Alice was excellent at remembering and analyzing other people's ideas, she was not very good at coming up with ideas of her own. Consequently she faltered in advanced schooling, where it is necessary (as it is in much of life) to have original ideas and not just to remember or analyze what one's teacher has said.

As an adult, Alice will face challenges if she does not develop higher levels of creative and practical skills. If, for example, she takes a job in finance, she will probably do well as an analyst but then be in trouble if she makes it to the next level and has to deal with high levels of uncertainty and change. If she goes into academia, she will do well in analyzing and critiquing the works of others but have difficulty coming up with her own original ideas or syntheses. If she tries law, she will likely succeed as an associate, but her lack of practical skills may cause her to have problems in the courtroom or even in working with clients. Thus one hopes that she will indeed develop a broader repertoire of skills.

Consider, in contrast, Barbara, a highly creative student. Barbara's grades were good although by no means spectacular. Her teachers thought she was just terrific as a source of new and exciting ideas despite the fact that her standardized test scores were very weak. She also did not excel in courses that mostly required memorization. Despite Barbara's mediocre scores on standardized tests and in some courses, she was enormously creative. As an undergraduate, she was publishing articles in refereed journals. If one were to look at her in a standard way—in terms of her ability to memorize—she would look third-rate at best. If one were to count her creativity, she might well appear as one of the best students in her class.

Barbara's creativity in itself will get her far, but perhaps not as far as she would wish. But creative skills in themselves are not enough. If she goes into design—art, architecture, advertising, or whatever—she will need somehow to develop skill in distinguishing her better ideas from her not-so-good ones. If she does sales, she will need to hone her practical skills so that she comes up not only with creative ways of selling products but also of connecting with other people. In general, people who are creative but not practical are often frustrated because they cannot convince others to adopt their great ideas.

Celia was a student whose academic performance was good but not great. She had good academic skills but none that made her stand out. She gave us quite a surprise when it came to getting a job. Everyone wanted to hire her. And that raised an intriguing question. Why would someone who lacked Alice's analytical ability and Barbara's creative ability do spectacularly well in the job market? The answer was that Celia had an abundance of practical intelligence or, put simply, common sense. Celia could go into an environment, figure out what she needed to do to adapt successfully in that environment, and then do it. For example, Celia knew how to interview effectively, how to interact well with other students, how to get her work done. She was also aware of the kinds of things that do and do not work in an academic environment. She knew something that is seldom acknowledged: that in school as in other aspects of life, one needs a certain amount of practical savvy in order to succeed.

Celia's practical abilities are a great asset that will be useful to her in later life. But she will need other kinds of skills as well. People with high levels of practical skills without corresponding creative skills usually end up selling ideas, but those of others rather than of their own. Moreover, if they lack analytical skills, they may find themselves selling ideas that are not very good but doing so successfully because of their ability to connect with people. In politicians, where I suspect such a pattern is common, there often seems to be a wide gap between the individual's ideas and his or her ability to sell them, with the latter outshining the former.

It is interesting to compare Alice, Barbara, and Celia with Paul. Paul combined Alice's analytical skills with Barbara's creative ones. As a result, many faculty members thought he would be extremely successful. I did not. The reason was that although Paul was very bright in some ways, he was notably challenged with regard to practical intelligence. He was the kind of student who was bright and who knew it. He had become very arrogant. But even arrogant people can get where they want to go if they know how to control their arrogance. Paul did not know. Although he received many job interviews, he was offered only one job, the worst one for which he applied. The reason was that he was unable or unwilling to hide his arrogance the one day on which it is essential to do so—the day of the job interview. Almost no one wanted to hire him, and he did not last long even at the place that finally did hire him.

I have described these four students at some length in order to argue that intelligence is not merely what intelligence tests test. These skills matter, particularly in courses that emphasize memory and analysis. Intelligence tests, but also other tests of cognitive and academic skills, thus 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. Practical and academic (or analytical) intelligence are separable sets of skills with different consequences. Even the tests that teachers use to measure achievement in college may be overly narrow if they do not assess students' skills in going well beyond the knowledge with which they have been presented (creative thinking) and students' skills in applying that knowledge in the real world (practical thinking).

Permit me to add one last example. As a freshman at Yale, I was extremely eager to major in psychology because I had done poorly on IQ tests as a child and wanted to understand why. I took the introductory psychology course and got a C. My professor at one point stared at me and commented that there was a famous Sternberg in the field of psychology (Saul Sternberg) and that it was obvious there would not be another one. Thirty-five years later, I was back in the same institution as a chaired professor and as president of the American Psychological Association. I commented to my predecessor, Phil Zimbardo, a professor at Stanford, that it was ironic that the president of the association, the largest association of psychologists in the world, was a C student in introductory psychology. At that point, he commented that he, too, had received a C in this course. This experience clarified for me the great difference between the memorization and rudimentary analytical skills required for success in many college courses, especially introductory ones, and the skills needed for success in later-life careers. I had never once had to memorize a book or lecture as a professional but I had had to teach students individually and in groups, do research, get grant proposals funded, negotiate the prickly world of academic politics, and so forth. I had needed a blend of analytical, creative, and practical skills that were only minimally tapped by the teaching and assessment of my early college years, especially in my chosen field.

The problem is that many students may be discouraged from pursuing careers in which they can be successful and contribute greatly because they do not do well in the introductory courses that often serve as gating mechanisms for determining who will go on to the advanced courses. If the skills required for success in these introductory courses are largely different from the skills required for success in more advanced studies and especially later in a corresponding career, potentially successful people may drop out of the field. Moreover, those who stay in the field may not be those who will succeed best later. Probably all of us who teach graduate students have encountered at least a handful who are good at taking knowledge-based tests but who lack the creative, analytical, and practical skills they need to do research and to practice in a given field.

Academic or analytical intelligence is the primary vehicle for success through much of college unless, in advanced (or even intermediate) courses, students are encouraged to produce projects, products, and portfolios that require more creative and practical thinking. Typically, the role of creative and practical intelligence becomes more important toward the end of college. But in schools with large classes, even in the later years, the creative and practical aspects may never come to the fore.

The Academic/Analytic Versus the Creative and Practical Aspects of Intelligence

The discussion so far suggests that intelligence is broader than the notion of general intelligence (*g*) that has dominated discussions in the past century (Carroll, 1993; Jensen, 1998; Spearman, 1904). One might wonder why there is a need to distinguish between analytical (academic) intelligence on the one hand and creative and practical intelligence on the other. Consider some of the data.

Creative Intelligence and General Cognitive Ability

When we think of creativity, eminent artists or scientists such as Michelangelo or Einstein immediately come to mind. However, these highly creative people are quite rare and difficult to study in the psychological laboratory. In his American Psychological Association address, Guilford (1950) noted that these problems had resulted in limited research on creativity. He proposed that creativity could be studied in everyday subjects using paper-and-pencil tasks. One of these was the Unusual Uses Test, in which an examinee thinks of as many uses for a common object (e.g., a brick) as possible. Many researchers adopted Guilford's suggestion, and "divergent thinking" tasks quickly became the main instruments for measuring creative thinking. The tests were a convenient way of comparing people on a standard "creativity" scale.

Building on Guilford's work, Torrance (1974) developed the Torrance Tests of Creative Thinking. These tests consist of several relatively simple verbal and figural tasks that involve divergent thinking plus other problem-solving skills. The tests can be scored for fluency (total number of relevant responses), flexibility (number of different categories of relevant responses), originality (the statistical rarity of the responses), and elaboration (amount of detail in the responses). Some subtests from the Torrance battery include the following:

1. Asking questions: The examinee writes out all the questions he or she can think of based on a drawing of a scene.
2. Product improvement: The examinee lists ways in which to change a toy monkey so children will have more fun playing with it.
3. Unusual uses: The examinee lists interesting and unusual uses of a cardboard box.
4. Circles: The examinee expands empty circles into different drawings and titles them.

A number of investigators have studied the relationship between creativity and intelligence, at least as measured by IQ. Three basic findings concerning creativity and conventional conceptions of intelligence are generally agreed upon (see, e.g., Barron & Harrington, 1981; Lubart, 1994).

First, creative people tend to show above-average IQs, often above 120 (see Renzulli, 1986). This figure is not a cutoff but rather an expression of the fact that people with low or even average IQs do not seem to be well represented among the ranks of highly creative individuals. Cox's (1926) geniuses had an estimated average IQ of 165. Barron estimated the mean IQ of his creative writers to be 140 or higher, based on their scores on the Terman Concept Mastery Test (Barron, 1963, p. 242). It should be noted that the Concept Mastery Test is exclusively verbal and thus provides a somewhat skewed estimate of IQ. The other groups in the Institute for Personality Assessment (IPAR) studies, that is, mathematicians and research scientists, were also above average in intelligence. Roe (1952, 1972), who did similarly thorough assessments of eminent scientists before the IPAR group was set up, estimated IQs for her participants that ranged between 121 and 194, with medians between 137 and 166, depending on whether the IQ test was verbal, spatial, or mathematical.

Second, above an IQ of 120, IQ does not seem to matter as much to creativity as it does below 120. In other words, creativity may be more highly correlated with IQ below an IQ of 120 but only weakly or not at all above an IQ of 120. (This relationship is often called the threshold theory.) In the architects' study, in which the average IQ was 130 (significantly above average), the correlation between intelligence and creativity was −.08, not significantly different from zero (Barron, 1969, p. 42). However, in the military officers' study, in which participants were of average intelligence, the correlation was .33 (Barron, 1963, p. 219). These results suggest that extremely highly creative people often have high IQs, but not necessarily that people with high IQs tend to be extremely creative (see also Getzels & Jackson, 1962). Some investigators (e.g., Simonton, 1994; Sternberg, 1996) have suggested that a very high IQ may actually interfere with creativity. Those who have very high IQs may be so highly rewarded for their IQ-like (analytical) skills that they fail to develop the creative potential within them, which may then remain latent.

Third, the correlation between IQ and creativity is variable, usually ranging from weak to moderate (Flescher, 1963; Getzels & Jackson, 1962; Guilford, 1967; Herr, Moore, & Hasen, 1965; Torrance, 1962; Wallach & Kogan, 1965; Yamamoto, 1964). The correlation depends in part on what aspects of creativity and intelligence are being measured, how they are being measured, and in what field the creativity is manifested. The role of intelligence is different in art and music, for instance, than in mathematics and science (McNemar, 1964). An obvious drawback to the tests used and assessments done by Roe and Guilford is the time and expense involved in administering them as well as the subjective scoring of them.

In contrast, Mednick (1962) produced a 30-item objectively scored 40-minute test of creative ability called the Remote Associates Test (RAT). It is based on his theory that the creative thinking process is the "forming of associative elements into new combinations which either meet specified requirements or are in some way useful. The more mutually remote the elements of the new combination, the more creative the process or solution" (Mednick, 1962). Because the ability to make these combinations and arrive at a creative solution necessarily depends on the existence of the combinations (i.e., the associative elements) in a person's knowledge base and the probability and speed of attainment of a creative solution are influenced by the organization of the person's associations, Mednick's theory suggests that creativity and intelligence are closely related; they are overlapping sets. Moderate correlations of .55, .43, and .41 have been shown between the RAT and the WISC (Wechsler Intelligence Scale for Children), the verbal portion of the SAT (formerly called the Scholastic Aptitude Test), and the Lorge-Thorndike Verbal intelligence measures, respectively (Mednick & Andrews, 1967). Correlations with quantitative intelligence measures were lower (r = .20 – .34). Correlations with other measures of creative performance have been more variable (Andrews, 1975). This psychometric approach for measuring creativity had both positive and negative effects on the field. On the positive side, the tests facilitated research by providing a brief, easy-to-administer, objectively scorable assessment device. Furthermore, research was now possible with "everyday" people (i.e., noneminent samples). However, there were also some negative effects.

First, some researchers criticized brief paper-and-pencil tests as trivial, inadequate measures of creativity, saying that larger productions such as actual drawings or writing samples should be used instead. Second, other critics suggested that no fluency, flexibility, originality, and elaboration scores captured the concept of creativity. In fact, the definition and criteria for creativity are matters of ongoing debate, and relying on the objectively defined statistical rarity of a response with regard to all the responses of a subject population is only one of many options. Other possibilities include using the social consensus of judges (see Amabile, 1983). Third, some researchers were less enchanted by the assumption that noneminent samples could shed light on eminent levels of creativity, which was the ultimate goal for many studies of creativity (e.g., Simonton, 1984). Thus a certain malaise developed, and this continues to accompany the paper-and-pencil assessment of creativity. Some psychologists, at least, have avoided this measurement quagmire in favor of less problematic research topics.

Practical Intelligence and General Cognitive Ability

General cognitive ability (g) is considered by many to be the best single predictor of job performance (e.g., Hunter, 1986; Ree & Earles, 1993; Ree, Earles, & Teachout, 1994; Schmidt & Hunter, 1998). The relationship between *g* and performance is attributed largely to the direct influence of *g* on the acquisition of job-related knowledge (Borman, 1991; Borman, Hanson, Oppler, & Pulakos, 1993; Hunter, 1986; Schmidt, Hunter, & Outerbridge, 1986). Many job-knowledge tests, however, assess primarily declarative knowledge of facts and rules (McCloy, Campbell, & Cudneck, 1994). They consist of abstract, well-defined problems that are similar to the types of problems found on traditional intelligence tests, thus explaining the observed correlations between measures of job knowledge and cognitive ability tests. Tests of practical intelligence, however, consist of problems that are poorly defined and context-specific. Such tests present practical, often work-related problems that test takers need to solve. We consider performance on these tests to be a function of practical rather than of abstract, general intelligence.

It is possible, however, to test practical intelligence as a separate entity. Tests of this construct usually give people situation-based problems related to their life or their work and ask them to resolve the problems. For example, a college student might be asked how to resolve a conflict with another student over the cleanliness of a dormitory room, or a business executive might be asked how he or she would resolve a production problem that has led to a lack of inventory.

Tests of practical intelligence exhibit trivial to moderate correlations with measures of *g*. In other words, they measure skills different from *g*. Scores on these tests for academic psychologists and for managers correlated nonsignificantly (–.04 to .16) with a test of verbal reasoning in undergraduate samples (Wagner, 1987; Wagner & Sternberg, 1985). Scores on a practical intelligence test for managers also exhibited a nonsignificant correlation with an IQ test for a sample of business executives (Wagner & Sternberg, 1990). Similar findings were obtained with a test of practical intelligence for sales in samples of undergraduates and salespeople (Wagner, Sujan, Sujan, Rashotte, & Sternberg, 1999). In one study, conducted in Kenya, practical intelligence scores actually correlated negatively with scores on tests of *g*, suggesting that, in certain environments, the development of practical skills may be emphasized at the expense of the development of academic skills (Sternberg et al., 2001). Such environments are not limited to rural Kenya: Artists, musicians, athletes, and craftsmen all may decide that the development of skills other than those taught in school may hold more value to them than do the more academic skills.

In a study by Eddy (1988), the Armed Services Vocational Aptitude Battery (ASVAB) was administered along with a practical intelligence test for managers to a sample of Air Force recruits. The ASVAB is a multiple-aptitude battery measuring verbal, quantitative, and mechanical abilities, and has been found to correlate highly with other conventional cognitive ability tests. Scores on the practical intelligence test exhibited near-zero correlations with factor scores on the ASVAB, again suggesting a distinction between academic and practical intelligence. In research with military leaders, leaders at three levels of command completed Terman's (1950) Concept Mastery Test, a test of verbal reasoning, along with a practical intelligence test for their respective levels. Practical intelligence scores exhibited trivial and nonsignificant to moderate and significant correlations (.02 to .25) with verbal reasoning ability (Hedlund et al., 2003).

The research reviewed above supports the contention that practical intelligence tests measure abilities that are distinct from those assessed by traditional intelligence tests. But how do scores on these practical tests relate to job performance?

Practical Intelligence and Performance

First, consider job knowledge. In terms of predicting job performance, tests of job knowledge have been found to predict performance fairly consistently, with an average validity of .48 after correcting for various statistical factors (Schmidt & Hunter, 1998). Much of this prediction is attributed to the relationship between job knowledge and general cognitive ability tests (Borman et al., 1993; Hunter, 1986). In other words, people with high *g* are expected to gain more knowledge and thus perform more effectively.

Practical intelligence tests have also been found to predict performance in a number of domains, typically correlating generally in the range of .2 to .5 with criteria such as rated prestige of business or institution, salary, performance appraisal ratings, number of publications, grades in school, and adjustment to college (Sternberg et al., 2000; Sternberg, Wagner, Williams, & Horvath, 1995; Wagner, 1987; Wagner & Sternberg, 1985). Below, we review some of these findings in more detail.

In studies with general business managers, practical intelligence scores correlated in the range of .2 to .4 with criteria such as salary, years of management experience, and whether or not the manager worked for a company at the top of the Fortune 500 list (Wagner, 1987; Wagner &

Sternberg, 1985). In a study with bank managers, Wagner and Sternberg (1985) obtained significant correlations between practical intelligence scores and average percentage of merit-based salary increase ($r = .48$, $p < .05$) and average performance rating for the category of generating new business for the bank ($r = .56$, $p < .05$).

Although much of the practical intelligence research has involved business managers (Sternberg et al., 2000), there is evidence that practical intelligence explains performance in other domains. In the field of academic psychology, correlations in the range of .3 to .4 were found between practical intelligence test scores and criterion measures such as citation rate, number of publications, and quality of department (Wagner, 1987; Wagner & Sternberg, 1985). Wagner and colleagues (1999) found correlations in the range of .3 to .4 between the tacit knowledge of salespeople and criteria such as sales volume and sales awards received.

Two studies showed the incremental validity of practical intelligence tests over traditional intelligence tests in predicting performance. In a study with business executives attending a Leadership Development Program at the Center for Creative Leadership, Wagner and Sternberg (1990) obtained a correlation of .61 between scores on a practical intelligence test for managers and performance on a managerial simulation. Furthermore, practical intelligence scores explained 32% of the variance in performance beyond scores on a traditional IQ test and also explained variance beyond measures of personality and cognitive style. In a study with military leaders, Hedlund et al. (2003) found practical intelligence scores to correlate significantly at all three levels of command with ratings of leadership effectiveness made by subordinates, peers, or superiors, with correlations ranging from .14 to .42 (Hedlund et al., 2003). More importantly, practical intelligence scores accounted for small (4%–6%), but significant variance in leadership effectiveness beyond scores on tests of general verbal intelligence and tacit knowledge for managers.

In a study with Yup'ik Eskimo children, Grigorenko et al. (2004) found that practical intelligence was a better predictor of hunting, fishing, gathering, and related adaptive skills than was academic intelligence. Moreover, whereas urban children outperformed rural Yup'ik children on conventional intelligence tests, the rural Eskimo children outperformed the urban children on tests relevant to the adaptive demands of their rural lives.

Other researchers, using practical intelligence tests or similar measures, also have found support for the relationship between practical intelligence and performance (e.g., Colonia-Willner, 1998; Fox & Spector, 2000; Pulakos, Schmitt, & Chan, 1996). Colonia-Willner administered the Tacit Knowledge Inventory for Managers (TKIM; Wagner & Sternberg, 1991), a test of practical intelligence, to bank managers along with measures of psychometric and verbal reasoning. She found that scores on the TKIM significantly predicted an index of managerial skill, whereas psychometric and verbal reasoning did not. Fox and Spector administered a situational judgment test to undergraduate students participating in a simulated interview. The students were asked to select the response they would most likely or least likely make to several work-related situations. Fox and Spector found that practical intelligence significantly predicted evaluations of the interviewee's qualifications. They also found that scores on the practical intelligence test exhibited a moderate, significant correlation (.25) with a measure of general intelligence. Finally, Pulakos et al. (1996), using a practical intelligence test specifically designed for entry-level professionals in a federal investigative agency, found that practical intelligence predicted both peer and supervisory ratings of performance. Furthermore, the effects of practical intelligence were not accounted for by *g*. Thus there is growing evidence to suggest that tests of practical intelligence and related tests not only explain individual differences in performance but also measure an aspect of performance that is not explained by measures of general intelligence.

In sum, if we were to consider only the academic aspect of intelligence, we would be considering only one aspect of what contributes to success in the world of work. If college were to develop only an academic skill set in students, it would ill prepare them for the world they would later confront.

CREATIVITY

Creativity is the potential to produce and implement ideas that are novel and high in quality. It goes beyond the creative intelligence discussed earlier in that it contains attitudinal, motivational, personality, and environmental components as well as the cognitive one of creative intelligence. Creativity is not an attribute limited to the historic "greats," such as Charles Darwin in the sciences, Pablo Picasso in art, Victor Hugo in literature, Bill Gates in technological entrepreneurship, or Abraham Lincoln in statesmanship. Rather, it is something anyone can use. In large part, it is a decision. So when we think about "creative ability," we think of it largely in terms of a person's decision to be creative—or not to be.

Creativity in a broad sense is important to active and engaged citizenship and to professional success. In a world that is rapidly changing, the ability flexibly to handle new challenges is essential. In 2008, two investment banks failed in rapid succession: Bear-Stearns and then Lehman Brothers. Other investment banks, such as Goldman-Sachs, weathered the storm. What was the difference? It was the ability of top management to see that the future would not resemble the past and that investment vehicles that in the past had been successful (such as high-risk mortgage-based securities) no longer would be. Those who were stuck in the past paid in the present.

Amabile (1983, 1996; Collins & Amabile, 1999) has described creativity as the confluence of intrinsic motivation, domain-relevant knowledge and abilities, and creativity-relevant skills. The creativity-relevant skills include (a) a cognitive style that involves coping with complexities and breaking one's mental set during problem solving; (b) knowledge of heuristics for generating novel ideas, such as trying a counterintuitive approach; and (c) a work style characterized by concentrated effort, an ability to set aside problems, and high energy.

Gruber and colleagues (Gruber & Davis, 1988) have proposed a developmental *evolving-systems model* for understanding creativity. A person's knowledge, purpose, and affect grow over time, amplify deviations that are encountered, and lead to creative products. Developmental changes in the knowledge system have been documented in cases such as Charles Darwin's thoughts on evolution. *Purpose* refers to a set of interrelated goals, which also develop and guide an individual's behavior. Finally, the affect or mood system notes the influence of joy or frustration on the projects undertaken.

Csikszentmihalyi (1988) has taken a different "systems" approach and highlights the interaction of the individual, domain, and field. An individual draws upon information in a domain and transforms or extends it via cognitive processes, personality traits, and motivation. The field, consisting of people who control or influence a domain (e.g., art critics and gallery owners), evaluates and selects new ideas. The domain, a culturally defined symbol system, preserves and transmits creative products to other individuals and future generations.

Gardner (1993a) has conducted case studies suggesting that the development of creative projects may stem from an anomaly within a system (e.g., tension between competing critics in a field) or moderate asynchronies between the individual, domain, and field (e.g., unusual individual talent for a domain). In particular, Gardner (1993a) has analyzed the lives of seven individuals who made highly creative contributions in the twentieth century, with each specializing in one of the multiple intelligences (Gardner, 1983b).

Although creativity can be understood in terms of uses of the multiple intelligences to generate new and even revolutionary ideas, Gardner's (1993b) analysis goes well beyond the intellectual. For example, Gardner pointed out two major themes in the behavior of these creative giants. First, they tended to have a matrix of support at the time of their creative breakthroughs. Second, they tended to drive a "Faustian bargain" whereby they gave up many of the pleasures people typically enjoy in life in order to attain extraordinary success in their careers. It is not clear that these attributes are intrinsic to creativity per se, however; rather, they seem to be associated with those who have been driven to exploit their creative gifts in a way that leads them to attain eminence.

Gardner further followed Csikszentmihalyi (1988, 1996) in distinguishing between the importance of the domain (the body of knowledge about a particular subject area) and the field (the context in which this body of knowledge is studied and elaborated, including the persons working with the domain, such as critics, publishers, and other "gatekeepers"). Both are important to the development and ultimately the recognition of creativity.

Investment Theory of Creativity

According to the investment theory of creativity, upon which this essay draws, creative thinkers are like good investors: they buy low and sell high (Sternberg & Lubart, 1995, 1996). They are willing and able to defy the crowd. Note that creative intelligence (ability) is important to creativity, but so is the attitude that defying the crowd is worthwhile and the motivation actually to defy the crowd. Whereas investors buy low and sell high in the world of finance, creative people do so in the world of ideas. Creative people generate ideas that are like undervalued stocks (stocks with a low price-to-earnings ratio), and both the stocks and the ideas are generally rejected by the public. When creative ideas are proposed, they often are viewed as bizarre, useless, and even foolish and are summarily rejected. The person proposing them is often regarded with suspicion and perhaps even with disdain and derision.

Creative ideas are both novel and valuable. But they are often rejected because the creative innovator stands up to vested interests and defies the crowd. The crowd does not maliciously or willfully reject creative notions. Rather, it does not realize and often does not want to realize that the proposed idea represents a valid and advanced way of thinking. Society generally perceives opposition to the status quo as annoying and offensive. This perception is seen as reason enough to ignore innovative ideas.

Evidence abounds that creative ideas are often rejected (Sternberg & Lubart, 1995). Initial reviews of major works of literature and art are often negative. Toni Morrison's *Tar Baby* received negative reviews when it was first published, as did Sylvia Plath's *The Bell Jar.* The first exhibition in Munich of the work of Norwegian painter Edvard Munch opened and closed on the same day because of the strong negative response from the critics. Some of the greatest scientific papers have been rejected not just by one journal but even by several journals before being published.

From the investment view, then, the creative person buys low by presenting a unique idea and then attempting to convince other people of its value. After convincing others that the idea is valuable, which increases the perceived value of the investment, the creative person sells high by leaving the idea to others and moving on to another idea. People typically want others to love their ideas, but immediate universal applause for an idea usually indicates that it is not particularly creative.

Creativity is as much a decision about and an attitude toward life as it is a matter of ability. Creativity is often obvious in young children but harder to find in older children and adults because their creative potential has been suppressed by a society that encourages intellectual conformity.

Creative work, and the broad-based creativity underlying it, requires applying and balancing the three intellectual abilities—creative, analytical, and practical—all of which can be developed (Sternberg & Grigorenko, 2007; Sternberg & Williams, 1996). Creative ability is used to generate ideas. Everyone, even the most creative person, has better and worse ideas. Without well-developed analytical ability, the creative thinker is as likely to pursue bad ideas as to pursue good ones. The creative individual uses analytical ability to work out the implications of a creative idea and to test it. Practical ability is used to translate theory into practice and abstract ideas into practical accomplishments. It is also used to convince other people that an idea is valuable. For example, every organization has a set of ideas that dictate how things, or at least some things, should be done. When an individual proposes a new procedure, he or she must sell it by convincing others that it is better than the old one. Practical ability is also used to recognize ideas that have a potential audience.

Creativity requires these three skills. The person who is only creatively intelligent may come up with innovative ideas but not be able to recognize or sell them. The person who is only analytical may be an excellent critic of other people's ideas but is not likely to generate creative ideas. The person who is only practical may be an excellent implementer or salesperson but is as likely to implement or promote ideas or products of little value as to promote genuinely creative ideas.

What are some particular characteristics one can seek to develop in an individual so that he or she decides for creativity? Put another way, what kinds of attributes are important for creative thought and action? They include (a) redefining problems; (b) questioning and analyzing assumptions; (c) selling creative ideas with the realization that the ideas will not sell themselves; (d) recognizing that knowledge is a double-edged sword—it can help or hurt creativity; (e) willingness to surmount obstacles; (f) willingness to take sensible risks; (g) tolerance of ambiguity; (h) self-efficacy; (i) finding what one loves to do; (j) willingness to delay gratification; (k) having a sense of humor and a sense of perspective about one's own strengths and limitations; and (l) having the courage to defy the crowd.

Creativity, like intelligence, is essential for effective citizenship and for making a difference to the world. But people can be intelligent and even creative but also foolish. Why? What attribute do they lack? I believe the attribute they lack is wisdom.

WISDOM

Wisdom may be the most important attribute to seek in future citizens and professionals. People can be academically or practically intelligent or creative but not wise. People who use their cognitive skills for evil or even selfish purposes or who ignore the well-being of others may be smart—but foolish.

Historically, the concept of wisdom has been the object of philosophical inquiries (Robinson, 1990) since the Platonic dialogues in *The Republic.* More recently, with the emergence of psychology as a field of study separate from philosophy, the concept of wisdom has also been studied as a psychological construct, and a number of psychologists have attempted empirical investigation of the concept of wisdom and its manifestations (Sternberg, 1990; Sternberg & Jordan, 2005). Wisdom has been studied from a range of psychological perspectives (Sternberg, 2001). Some researchers (see Clayton 1975, 1982; Holliday & Chandler, 1986; or Sternberg, 1990) have focused on implicit theories of wisdom—

that is, on trying to understand how the layperson perceives and defines wisdom. Other researchers have adopted a developmental perspective to investigate how wisdom develops or fails to develop. Most noticeably, empirical work in this area was conducted by the late Paul Baltes and his colleagues at the Max Planck Institute in Berlin (e.g., Baltes & Staudinger, 1993, 2000). Another developmental approach to defining wisdom is to view it as postformal–operational thinking, extending beyond the traditional Piagetian stages of intelligence (Piaget, 1972).

Several researchers and theoreticians have focused on the importance of integration and balance in wisdom. Gisela Labouvie-Vief (1990), for example, has emphasized the balance between different kinds of thinking, suggesting that wisdom constitutes a balance of *logos,* which are objective and logical processes, and *mythos,* which represent subjective and organismic processes. Deborah Kramer (1990) has focused on the balance between various self-systems such as the cognitive, conative, and affective, arguing that wisdom involves integration of cognition and affect, resulting in a well-balanced personality, where the conscious and unconscious interact in harmony. Still others insist on the balance between different points of view (Kitchener & Brenner, 1990) or on "a balance between the opposing valences of intense emotion and detachment, action and inaction, knowledge and doubts" (Birren & Fisher, 1990, p. 326). The following pages focus on a theory of wisdom proposed by Sternberg, which builds on previous theories emphasizing the importance of integration and balance in wisdom.

The Balance Theory of Wisdom

According to Sternberg's balance theory of wisdom (Sternberg, 1998b, 2001), wisdom is the application of intelligence, creativity, and knowledge as mediated by positive ethical values toward the achievement of a common good through a balance among (a) intrapersonal, (b) interpersonal, and (c) extrapersonal interests over the short and long term.

Wisdom is not just about maximizing one's own or someone else's self-interest but about balancing various self-interests (intrapersonal) with the interests of others (interpersonal) and of other aspects of the context in which one lives, such as one's city or country or environment or even God. Wise people—such as Nelson Mandela, Mother Teresa, or Martin Luther King—see far beyond their own personal interests to the interests of others and of society as a whole.

A person can be practically intelligent but use his or her practical intelligence toward bad or selfish ends. In wisdom, one certainly may seek good ends for oneself, but one also seeks the common good and good outcomes for others. If one's motivations are to maximize certain people's interests and minimize other people's, wisdom is not involved. In wisdom, one seeks a common good, realizing that this common good may be better for some than for others.

Problems requiring wisdom always involve at least some element of each of intrapersonal, interpersonal, and extrapersonal interests. For example, when a president decides to go to war or a CEO decides to introduce a wholly new product line, or a university president decides to open a new school (e.g., a law school or a medical school), the consequences are large and affect many persons as well as institutions. And the decision always has to be made in the context of the whole range of available options. But wisdom can apply in smaller decisions as well, such as whether a couple should move to a new location where one spouse has been offered a new job when the other spouse is already happily employed in their present place of residence.

What kinds of considerations might be included under each of the three kinds of interests? Intrapersonal interests might include the desire to enhance one's popularity or prestige, to make more money, to learn more, to increase one's spiritual well-being, to increase one's power, and so forth. Interpersonal interests might be quite similar, except as they apply to other people rather than oneself. Extrapersonal interests might include contributing to the welfare of one's school, helping one's community, contributing to the well-being of one's country, or serving God, and so forth. Different people balance these interests in different ways. At one extreme, a malevolent dictator might emphasize his or her own personal power and wealth; at the other extreme, a saint might emphasize serving only others and God.

What are the characteristics of people who are analytically intelligent and perhaps even creative but foolish? I propose five characteristics, based on Sternberg (2002).

The first is egocentrism. Many smart people have been so highly rewarded in their lives that they lose sight of the interests of others. They start to act as though the whole world revolved around them. In doing so, they often set themselves up for downfalls, as happened to Dennis Kozlowski, formerly CEO of Tyco, who spent company money extravagantly on himself and his wife as though company resources were to be tapped as his own personal piggybank.

The second characteristic is a false sense of omniscience. Smart people typically know a lot. They get in trouble, however, when they start to think that they "know it all." They may have expertise in one area, but then start to fancy themselves experts in practically everything. At that point they become susceptible to remarkable downfalls because they act as experts in areas where they are not and can make disastrous mistakes in doing so. Colin Powell, typically known for his wisdom, some years ago gave a speech to the United Nations that was widely watched around the world. It described in detail the evidence for the weapons of mass destruction that Iraq was alleged to possess. Many people were impressed with the seeming near omniscience of U.S. intelligence agencies about the Iraqi weapons program. It later turned out that almost the entire speech was based on false intelligence.

The third characteristic is a false sense of omnipotence. Many smart people find themselves in positions of substantial power. Sometimes they lose sight of the limitations of their power, and start to act as though they were omnipotent. Several U.S. presidents as well as presidents of other countries have had this problem, leading their countries to disasters on the basis of personal whims. Many corporate chieftains have also started to think of themselves as omnipotent, unfortunately cooking the books of their corporations at will. Robert Mugabe (who is, as I write this essay, the "president" of Zimbabwe) seems to feel that there is no indignity too great to heap upon his own countrymen. Adolf Hitler seemed to believe that no one and nothing could stop him in his quest for world domination.

The fourth characteristic is a false sense of invulnerability. Such individuals not only think they that can do anything but also believe that they can get away with it. They believe that either they are too smart to be found out or, even if found out, they will escape any punishment for their misdeeds. The result is the kind of disasters the United States has seen in the recent Enron, Worldcom, and Arthur Andersen debacles. Mugabe, mentioned above, has been quoted as saying that only God can remove him from power. That is about as close to invulnerable as a dictator can feel, in today's world or any other.

Andrew Fastow, a graduate of the university in which I teach, seemed to believe that no one would ever be able to decipher the complex fraudulent schemes by which he enriched himself at the expense of Enron employees, customers, and shareholders.

The fifth characteristic is a false sense of ethical disengagement. One comes to believe that ethics apply to others but not to oneself. Those who are ethically disengaged insist on ethical behavior from all people except themselves. A number of television evangelists—such as Jimmy Swaggert, Jim Bakker, and Ted Haggard—preached to tens of thousands of people about the wages of sin while committing at gross levels the very sins against which they were preaching.

Academic intelligence, practical intelligence, creativity, and wisdom are important attributes of success in life. When one comes up with an idea, one must be creative to ensure that the idea is novel; analytical to ensure that the idea is good; practical to ensure the idea can be implemented and that people can be persuaded of its value; and wise to ensure that the idea helps achieve a common good.

For these characteristics to be useful as criteria in a program, they need to be identifiable. How does one identify these attributes in individuals? Can they be taught and measured?

TEACHING AND ASSESSMENT WITH WICS

The question arises: How can the WICS model be operationalized? Consider, in turn, teaching and assessment.

It is important to teach students not just to learn facts but also to think analytically, creatively, practically, and wisely (Sternberg & Grigorenko, 2007; Sternberg, Jarvin, & Grigorenko, 2009).

Teaching Analytically

Teaching *analytically* means encouraging students to (a) analyze, (b) critique, (c) judge, (d) compare and contrast, (e) evaluate, and (f) assess. When teachers refer to teaching for "critical thinking," they typically mean teaching for analytical thinking. How does such teaching translate into instructional and assessment activities? Consider the following examples across the school curriculum:

1. Analyze the development of the character of Heathcliff in *Wuthering Heights.* [Literature]
2. Critique the design of the experiment (just gone over in class or in a reading) showing that certain plants grow better in dim light than in bright sunlight. [Biology]
3. Judge the artistic merits of Roy Lichtenstein's "comic-book art," discussing its strengths as well as its weaknesses as fine art. [History of Art]
4. Compare and contrast the respective natures of the American Revolution and the French Revolution, pointing out ways both in

which they were similar and those in which they were different. [History]

5. Evaluate the validity of a solution to a mathematical problem and discuss any weaknesses in the solution. [Mathematics]
6. Assess the strategy used by a winning player in a tennis match you just observed, stating what techniques she used in order to defeat her opponent. [Physical Education]

Teaching Creatively

Teaching *creatively* means encouraging students to (a) create, (b) invent, (c) discover, (d) imagine if . . . , (e) suppose that . . . , (f) predict. Teaching for creativity requires teachers not only to support and encourage creativity but also to role-model it and to reward it when it is displayed (Sternberg & Lubart, 1995; Sternberg & Williams, 1996). In other words, teachers need not only to talk the talk but also to walk the walk. Consider some examples of instructional or assessment activities that encourage students to think creatively:

1. Create an alternative ending to a short story you just read that represents a different way things might have gone for the main characters in the story. [Literature]
2. Invent a dialogue between an American tourist in Paris and a Frenchman he encounters on the street from whom he is asking directions on how to get to the Rue Pigalle. [French]
3. Discover the fundamental physical principle that underlies all of the following problems, each of which differs from the others in the “surface structure” of the problem but not in its “deep structure . . . .” [Physics]
4. Imagine that the government of China keeps evolving over the course of the next 20 years in much the same way as it has been evolving. What do you believe the government of China will be like in 20 years? [Government/Political Science]
5. Suppose that you were to design one additional instrument to be played in a symphony orchestra for future compositions. What might that instrument be like and why? [Music]
6. Predict changes that are likely to occur in the vocabulary or grammar of spoken Spanish in the border areas of the Rio Grande over the next 100 years as a result of continuous interactions between Spanish and English speakers. [Linguistics]

Teaching Practically

Teaching *practically* means encouraging students to (a) apply, (b) use, (c) put into practice, (d) implement, (e) employ, and (f) render practical what they know. Such teaching must relate to the real practical needs of the students, not just to what would be practical for individuals other than the students. Consider some examples:

1. Apply the formula for computing compound interest to a problem people are likely to face when planning for retirement. [Economics, Math]
2. Use your knowledge of German to greet a new acquaintance in Berlin. [German]
3. Put into practice what you have learned from teamwork in football to making a classroom team project succeed. [Athletics]
4. Implement a business plan you have written in a simulated business environment. [Business]
5. Employ the formula for distance, rate, and time to compute a distance. [Math]
6. Render practical a proposed design for a new building that will not work in the esthetic context of the surrounding buildings, all of which are at least 100 years old. [Architecture]

Teaching Analytically, Creatively, and Practically

These examples emphasize how it is possible to teach analytical, creative, and practical thinking as distinct sets of skills. In real life, however, the large majority of problems require some combination of the three sets of skills. Consider, for example, the problem of a business, such as a technology company, that realizes its product lines are becoming obsolete. Once upon a time, it might have created buggy whips. Today, it might create SUVs. What should the company do? It might decide to close or go into bankruptcy. But, in many cases companies can survive if they define the problem in a different way. For example, the company specializing in SUVs might start emphasizing smaller, more fuel-efficient cars. Or it might switch to smaller SUVs. Or it might merge with a company that produces smaller cars. It needs creative managers to come up with ideas to save it. It needs analytical managers to evaluate the various options. And it needs practical managers who can take the ideas and make them work in practice. But it also needs wise managers who think about larger questions, such as whether there is still a role for SUVs in a society in which gas prices are soaring and large vehicles are polluting the world to a dangerous extent.

Teaching for Wisdom

Teachers who teach for *wisdom* will explore with students the notion that conventional abilities and achievements are not enough for professional success as well as a meaningful life. Many people become trapped in their lives and, despite feeling conventionally successful, feel that their lives lack meaning, at least in the sense of contributing to a good that is larger than their own. Contributing to a common good is not an alternative to success but rather an aspect of it that, for most people, goes beyond money, promotions, large houses, and so forth. The teacher will further demonstrate how wisdom is critical for a life that makes a positive difference to the world. In the long run, wise decisions benefit people in ways that foolish decisions never do. The teacher must teach students the usefulness of interdependence—a rising tide raises all ships, while a falling tide can sink them.

It is also important to be a role model of wisdom, because what you do is more important than what you say. Wisdom is in what you do, not just in what you say. So students should read about wise judgments and decision making in the context of the actions that followed, so that the students understand that such means of judging and decision making exist. Teachers need to help students to learn to recognize their own interests, those of other people, and those of institutions. They need further to help students learn to balance their own interests, those of other people, and those of institutions. They will teach students that the "means" by which the end is obtained matters, not just the end. Students need to be encouraged to form, critique, and integrate their own values in their thinking. They further need to learn to think dialectically, realizing that both questions and their answers evolve over time, and that the answer to an important life question can differ at different times in one's life (such as whether to marry). Wisdom further requires them to learn to think dialogically, whereby they understand interests and ideas from multiple points of view. For example, what one group views as a "settler," another may view as an "invader." Most importantly, students need to learn to search for and then try to reach the common good—a good where everyone wins, not only those with whom one identifies.

Teaching for wisdom will succeed only if teachers encourage and reward wisdom. Teachers must make wisdom real for students' lives. Teachers should teach students to monitor events in their lives and their own thought processes about these events. One way to learn to recognize others' interests is to begin to identify your own. Teachers should also help students to understand the importance of inoculating oneself against the pressures of unbalanced self-interest and small-group interest.

Students will develop wisdom by becoming engaged in class discussions, projects, and essays that encourage them to discuss the lessons they have learned from both classical and modern works and how these lessons can be applied to their own lives and the lives of others. They need to study not only "truth" as we know it but also values. The idea is not to force-feed a set of values but to encourage students reflectively to develop their own prosocial ones.

Students should be encouraged to think about how almost everything they study might be used for better or worse ends and to realize that the ends to which knowledge is put *do* matter. Teachers need to realize that the only way they can develop wisdom in their students is to serve as role models of wisdom themselves. A role model of wisdom will, I believe, take a much more Socratic approach to teaching than teachers customarily do. Students often want large quantities of information spoon-fed or even force-fed to them. They then attempt to memorize this material for exams, only to forget it soon thereafter. In a wisdom-based approach to teaching, students will need to take a more active role in constructing their learning.

For example, in history, one might ask whether the term *settler* has a different meaning to those who are settling versus those who view themselves as already living the land being settled (e.g., American Indians in the 18th century). In science, one might ask how a new discovery or invention could be put to good or base use for society. In literature, one might ask whether a great literary figure, such as King Lear, was wise, and if not, why not. In foreign language, one might ask whether there is wisdom to be learned from another culture that our culture seems to lack. In mathematics, one might ask how mathematical formulas, such as those forming the bases of building bridges, can create a better life for many people. In art, one might ask how wise versus foolish people are depicted by different artists. In music, one might ask whether one can appreciate the music of people such as Richard Wagner, who had hateful prejudices against certain groups (in Wagner's case, Jews).

Slavery is a topic that lends itself particularly well to teaching for wisdom (Reznitskaya & Sternberg, 2004; Sternberg, Reznitskaya, & Jarvin, 2007). How could a whole country allow slavery, as the United States once did? What were the people thinking? Could a country allow such thinking only by viewing the slaves as less than human? Did slavery show the human capacity for rationalization in the place of wisdom? In our own unit on slavery in a program for teaching American history, we consider the formation, implementation, and abolition of slavery from the standpoint of wisdom-based thinking (see Reznitskaya & Sternberg, 2004).

We are implementing many of the ideas described here in a curriculum that has been developed at Tufts University for a leadership minor. The leadership minor provides an opportunity for undergraduate students to learn about leadership as it applies in whatever discipline they happen to be studying. The minor has three tiers. The first comprises courses on leadership and ethics, such as Psychology of Leadership and Sociology of Leadership. The second tier comprises courses across many departments in the liberal arts and sciences that touch on leadership. For example, students might study leadership through the eyes of Macbeth or of an 18th-century artist drawing leaders, or of a historian studying the American Revolution, or of a scholar of comparative political systems. The third tier comprises a leadership practicum, which requires the student to engage in a practical leadership experience and then to write a reflection paper on it that relates what he or she has learned in the practicum to what was learned in the classroom.

How does a leadership minor develop WICS? Consider an example from my own course on leadership. On the first day of my class last spring, I went over the syllabus, which involved a lot of work. A student raised his hand and challenged me, saying the course was too much work and it was obvious that I was out of touch with the lives of students. I explained to the class that this challenge was quite embarrassing, as I was supposed to be the leader of the class and now my authority was being publicly challenged. I explained that, in leadership, the question is not whether your authority will be publicly challenged—it will be—but rather how you respond to such challenges. I told them that I was just not sure how to respond to the challenge and that I hoped they would help me. I divided them into three groups and asked each group to simulate what had happened and then also to simulate how I should respond, as well as the likely potential outcome. I then thanked the shill I had planted in the class to do this demonstration (an assistant in the dean's office), and he left. This exercise and others like it gave students an opportunity to think creatively in generating a response to the challenge, analytically in deciding whether the response was a good one, practically in implementing their response through a simulation, and wisely in ensuring that the response helped to achieve a common good.

Assessment With WICS

If we are going to teach for WICS, we must learn to assess for WICS. In general, it is important in education that the way we assess skills matches the way we teach them. How does one assess the various WICS abilities? Following are described two studies, the first of which did not involve assessing wisdom and the second of which did. The goal of both of these studies is to select and provide a basis for developing reflective, active citizens and professionals who will make a positive difference in their own lives and the world. They will be citizens prepared for the challenges of the world, not just of what may be sometimes narrowly focused classrooms.

In a recent study supported by the College Board (Sternberg & 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 multiple-choice tests, we used 3 additional measures of creative skills and 3 of practical skills, as outlined here.

Creative Skills

The three additional tests were as follows:

1. *Cartoons.* Participants were given five cartoons purchased from the archives of the *New Yorker,* but with the caption removed. The participant's task was to choose three cartoons and provide a caption for each. Two trained judges rated all the cartoons for cleverness, humor, and originality. A combined creativity score was formed by summing the individual ratings on each dimension.
2. *Written Stories.* Participants were asked to write two stories, spending about 15 minutes on each, choosing from the following titles: "A Fifth Chance," "2983," "Beyond the Edge," "The Octopus's Sneakers," "It's Moving Backwards," and "Not Enough Time." A

team of four judges was trained to rate the stories for originality, complexity, emotional evocativeness, and descriptiveness.

3. *Oral Stories.* Participants were presented with five sheets of paper, each containing a set of pictures linked by a common theme. For example, participants might receive a sheet of paper with images of a musical theme, a money theme, or a travel theme. The participant then chose one of the pages and was given 15 minutes to formulate a short story and dictate it into a cassette recorder. The dictation period was not to be more than 5 minutes long. The process was then repeated with another sheet of images so that each participant dictated a total of two oral stories. Six judges were trained to rate the stories for originality, complexity, emotional evocativeness, and descriptiveness.

Practical Skills

The additional tests were as follows:

1. *Everyday Situational Judgment Inventory (Movies).* This video-based inventory presents participants with seven brief vignettes that capture problems encountered in general, everyday life, such as determining what to do when one is asked to write a letter of recommendation for someone one does not know particularly well.
2. *Common Sense Questionnaire.* This written inventory presents participants with 15 vignettes that capture problems encountered in general business-related situations, such as managing tedious tasks or handling a competitive work situation.
3. *College Life Questionnaire.* This written inventory presents participants with 15 vignettes that capture problems encountered in general college-related situations, such as handling trips to the bursar's office or dealing with a difficult roommate.

We found that our tests significantly and substantially improved upon the validity of the SAT for predicting first-year college grades (Sternberg & the Rainbow Project Collaborators, 2006), doubling prediction over the SAT alone, and increasing prediction by 50% over SATs and high school GPA. 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 GPA. Stemler, Grigorenko, Jarvin, and Sternberg (2006) found that it is possible as well to reduce ethnic-group differences on certain Advanced Placement (AP) examinations by including creative and practical questions on them.

The Rainbow Project did not assess wisdom, but wisdom is important in producing active and successful citizens and professionals. Tufts University in Medford, Massachusetts, has strongly emphasized the role of active citizenship in education. It has put into practice some of the ideas from the Rainbow Project. In collaboration with Dean of Admissions Lee Coffin, we instituted Project Kaleidoscope, which represents an implementation of the ideas of Rainbow but goes beyond that project to include in its assessment the construct of wisdom (for more details, see Sternberg, 2007, 2008a).

On the 2006–2007 application for all of the more than 15,000 students applying to the arts, sciences, and engineering programs at Tufts, we placed questions designed to assess wisdom, analytical and practical intelligence, and creativity synthesized (WICS); (Sternberg, 2003). The program was continued for 2007–2008 and is being continued for 2008–2009, but the data reported here are for the first year, for which we have more nearly complete data.

WICS is a theory that extends the theory of successful intelligence on the basis of the notion that some people may be academically and even practically intelligent but unwise, as in the case of corporate scandals and numerous political scandals as well. The perpetrators were smart, well-educated, and foolish. The conception of wisdom used here is that of the balance theory of wisdom (Sternberg, 1998b), according to which wisdom is the application of intelligence, creativity, and knowledge for the common good, by balancing intrapersonal, interpersonal, and extrapersonal interests over the long and short terms through the infusion of positive ethical values.

The questions were optional in the first 2 years. Whereas the Rainbow Project was done as a separate high-stakes test administered with a proctor, the Kaleidoscope Project was done as a section of the Tufts-specific supplement to the Common Application (Sternberg, 2007). It just was not practical to administer a separate high-stakes test such as the Rainbow assessment for admission to one university. Moreover, the advantage of Kaleidoscope is that it got us away from the high-stakes testing situation in which students must answer complex questions in very short amounts of time under incredible pressure.

Students were encouraged to answer just a single question so as not overly to burden them. Tufts University competes for applications with many other universities, and if our application was substantially more burdensome than those of our competitor schools, it would put us at a real-world disadvantage in attracting applicants. In the theory of successful intelligence, successful intelligent individuals capitalize on strengths and compensate for or correct weaknesses. Our format gave students a chance to capitalize on a strength.

As examples of items, a creative question asked students to write stories with titles such as "The End of MTV" or "Confessions of a Middle-School Bully." Another creative question asked students what the world would be like if some historical event had come out differently, for example, if Rosa Parks had given up her seat on the bus. Yet another creative question, a nonverbal one, gave students an opportunity to design a new product or an advertisement for a new product. A practical question queried how students had persuaded friends of an unpopular idea they held. A wisdom-based question asked students how a passion they had could be applied toward a common good.

Analytical quality was assessed by the cogency, organization, logic, and balance of the essay. Creativity was assessed by looking for novelty, quality, and task appropriateness. Practicality was a matter of the utility with respect to human and material resources as well as persuasiveness. Wisdom was assessed by the extent to which the response represented the use of abilities and knowledge for a common good by balancing one's own, others', and institutional interests over the long and short terms through the infusion of positive ethical values.

Note that the goal is not to replace SAT and other traditional admissions measurements like GPAs and class rank with some new test. Rather, it is to reconceptualize applicants in terms of academic/analytical, creative, practical, and wisdom-based abilities using the essay as one but not the only source of information. For example, highly creative work submitted in a portfolio also could be entered into the creativity rating, or evidence of creativity through winning of prizes or awards. The essays were major sources of information, but if other information was available, the trained admissions officers used it.

We now have some results of our first year of implementation, and they are very promising. If sufficient evidence was available, applicants were evaluated for creative, practical, and wisdom-based skills as well as for academic (analytical) and personal qualities in general.

Among the applicants who were evaluated as being academically qualified for admission, approximately half completed an optional essay. Merely doing these essays had no meaningful effect on chances of admissions. However, the *quality* of the essays or other evidence of creative, practical, or wisdom-based abilities did have an effect. For those rated at "A" (the top rating) by a trained admission officer in any of these three categories, average rates of acceptance were roughly double those for applicants not getting an A. Because of the large number of essays (over 8,000), only one rater rated applicants except for a sample to ensure that interrater reliability was sufficient, which it was.

Many measures do not look like conventional standardized tests but have statistical properties that mimic them. We were therefore interested in what is called the "convergent-discriminant validation" of our measures—showing that they measure what they should measure and do not measure what they should not measure. The correlation of our measures with a rated academic composite that included SAT scores and high school GPA were modest but significant for creative, practical, and wise thinking. The correlations with a rating of quality of extracurricular participation and leadership were moderate for creative, practical, and wise thinking. Thus, the pattern of convergent–discriminant validation was what we had hoped for: the assessments correlated with what they were supposed to correlate with and did not correlate with what they were not supposed to correlate with.

The average academic quality of applicants in arts and sciences rose slightly in 2006–2007, the first year of the pilot, in terms of both SAT and high school GPA. In addition, there were notably fewer students in what before had been the bottom third of the pool in terms of academic quality. Many of those students, seeing the new application, seem to have decided not to bother to apply. Many more strong applicants applied. After one semester of study, students who had A's on Kaleidoscope were performing just as well as other students.

Thus, adopting these new methods does not result in less qualified applicants applying to the institution and being admitted. Rather, the applicants who are admitted are *more* qualified, but in a broader way. Perhaps most rewarding were the positive comments from large numbers of applicants that they felt our application gave them a chance to show themselves for who they were. Of course, many factors are involved in admissions decisions, and Kaleidoscope ratings were only one small part of the overall picture.

We did not get meaningful differences across ethnic groups, a result that surprised us, given that the earlier Rainbow Project reduced but did not eliminate differences. And after a number of years in which applications by underrepresented minorities were relatively flat in terms of numbers, this year they went up substantially. In the end, applications from African Americans and Hispanic Americans increased significantly; admissions of African Americans were up 30% and of Hispanic Americans up 15%. So our results, like those of the Rainbow Project, showed that it is possible to increase academic quality and diversity simultaneously and to do so for an entire undergraduate class at a major university, not just for small samples of students at some scattered schools. Most importantly, we sent a message to students, parents, high school guidance counselors, and others that we believe that there is more to a person than the narrow spectrum of skills assessed by standardized tests, and that these broader skills can be assessed in a quantifiable way.

CONCLUSIONS

In developing the active, involved, and reflective citizens and professionals of tomorrow, some very important factors to consider are academic and practical intelligence, creativity, and wisdom—synthesized so that they work together effectively. I am not claiming that these are the only attributes that matter. For example, motivation is extremely important as well. I do believe, however, that motivation is partly (although not exclusively) situational, and that with the proper environment, anyone can be motivated to achieve.

A major problem we face is that the way we teach often does not correspond to the way students need to learn to succeed in life and particularly in the careers for which students are preparing. The skill sets required in school and on the job may overlap only weakly. The result may be unprepared job seekers and seekers of jobs who might be better off doing something other than what they have chosen. At the same time, students who might have chosen and succeeded in a career may be discouraged from entering that career because they did not succeed as well as they had hoped in the introductory courses supposed to prepare them for that career.

Colleges and universities 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. WICS 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 and educate students so as to maximize their positive future impact. We want our students to show wisdom. We must do the same.

REFERENCES

Amabile, T. M. (1983). *The social psychology of creativity.* New York: Springer-Verlag.

Amabile, T .M. (1996). *Creativity in context.* Boulder, CO: Westview.

Andrews, F. M. (1975). Social and psychological factors which influence the creative process. In I. A. Taylor & J. W. Getzels (Eds.), *Perspectives in creativity* (pp. 117–145). Chicago: Aldine.

Baltes, P. B., & Staudinger, U. (1993). The search for a psychology of wisdom. *Current Directions in Psychological Science, 2,* 75–80.

Baltes, P. B., & Staudinger, U. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist, 55,* 122–136.

Barron, F. (1963). *Creativity and psychological health.* Princeton, NJ: D. Van Nostrand.

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.

Birren, J. E., & Fisher, L. M. (1990). The elements of wisdom: Overview and integration. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 317–332). New York: Cambridge University Press.

Borman, W. C. (1991). Job behavior, performance, and effectiveness. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 2, pp. 271–326). Palo Alto, CA: Consulting Psychologists Press.

Borman, W. C., Hanson, M. A., Oppler, S. H., & Pulakos, E. D. (1993). Role of supervisory experience in supervisory performance. *Journal of Applied Psychology, 78,* 443–449.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Clayton, V. (1975). Erickson's theory of human development as it applies to the aged: Wisdom as contradictory cognition. *Human Development, 18,* 119–128.

Clayton, V. (1982). Wisdom and intelligence: The nature and function of knowledge in the later years. *International Journal of Aging and Development, 15,* 315–321.

Collins, M. A., & Amabile, T. M. (1999). Motivation and creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 297–312). New York: Cambridge University Press.

Colonia-Willner, R. (1998). Practical intelligence at work: Relationship between aging and cognitive efficiency among managers in a bank environment. *Psychology and Aging, 13,* 45–57.

Cox, C. M. (1926). *The early mental traits of three hundred geniuses.* Stanford, CA: Stanford University Press.

Csikszentmihalyi, M. (1988). Society, culture, and person: A systems view of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 325–339). New York: Cambridge University Press.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention.* New York: HarperCollins.

Eddy, A. S. (1988). The relationship between the Tacit Knowledge Inventory for Managers and the Armed Services Vocational Aptitude Battery. Unpublished master's thesis, St. Mary's University, San Antonio, TX.

Flescher, I. (1963). Anxiety and achievement of intellectually gifted and creatively gifted children. *Journal of Psychology, 56,* 251–268.

Fox, S., & Spector, P. E. (2000). Relations of emotional intelligence, practical intelligence, general intelligence, and trait affectivity with interview outcomes: It's not all just "G." *Journal of Organizational Behavior, 21,* 203–220.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Gardner, H. (1993a). *Creating minds: An anatomy of creativity seen through the lives of Freud, Einstein, Picasso, Stravinsky, Eliot, Graham, and Gandhi.* New York: HarperCollins.

Gardner, H. (1993b). *Multiple intelligences: The theory in practice.* New York: Basic Books.

Gardner, H. (1999). *Intelligence reframed.* New York: Basic Books.

Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students.* New York: Wiley.

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.

Gruber, H. E., & Davis, S. N. (1988). Inching our way up Mount Olympus: The evolving-systems approach to creative thinking. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 243–270). New York: Cambridge University Press.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5,* 444–454.

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

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.

Herr, E. L., Moore, G. D., & Hasen, J. S. (1965). Creativity, intelligence, and values: A study of relationships. *Exceptional Children, 32,* 114–115.

Holliday, S. G., & Chandler, M. J. (1986). *Wisdom: Explorations in adult competence.* Basel, Switzerland: Karger.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29,* 340–362.

Jensen, A. R. (1998). *The* g *factor.* Westport, CT: Praeger-Greenwood.

Kitchener, K. S., & Brenner, H. G. (1990). Wisdom and reflective judgment: Knowing in the face of uncertainty. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 212–229). New York: Cambridge University Press.

Kramer, D. A. (1990). Conceptualizing wisdom: The primacy of affect–cognition relations. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 279–313). New York: Cambridge University Press.

Labouvie-Vief, G. (1990). Wisdom as integrated thought: Historical and developmental perspectives. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 52–83). New York: Cambridge University Press.

Lubart, T. I. (1994). Creativity. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 290–332). San Diego, CA: Academic Press.

McCloy, R. A., Campbell, J. P., & Cudneck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology, 79,* 493–505.

McNemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist, 19,* 871–882.

Mednick, M. T., & Andrews, F. M. (1967). Creative thinking and level of intelligence. *Journal of Creative Behavior, 1,* 428–431.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 69,* 220–232.

Piaget, J. (1972). *The psychology of intelligence.* Totowa, NJ: Littlefield-Adams

Pulakos, E. D., Schmitt, N., & Chan, D. (1996). Models of job performance ratings: An examination of ratee race, ratee gender, and rater level effects. *Human Performance, 9,* 103–119.

Ree, M. J., & Earles, J. A. (1993). *g* is to psychology what carbon is to chemistry: A reply to Sternberg and Wagner, McClelland, and Calfee. *Current Directions in Psychological Science, 1,* 11–12.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79*, 518–524.

Renzulli, J. S. (1986). The three-ring conception of giftedness: a developmental model for creative productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 53–92). New York: Cambridge University Press.

Reznitskaya, A., & Sternberg, R. J. (2004). Teaching students to make wise judgments: The "teaching for wisdom" program. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 181–196). New York: Wiley.

Robinson, D. N. (1990). Wisdom through the ages. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 13–24). New York: Cambridge University Press

Roe, A. (1952). *The making of a scientist.* New York: Dodd, Mead.

Roe, A. (1972). Patterns of productivity of scientists. *Science, 176,* 940–941.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262–274.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71,* 432–439.

Simonton, D. K. (1984). *Genius, creativity, and leadership.* Cambridge, MA: Harvard University Press.

Simonton, D. K. (1994). *Greatness: Who makes history and why?* New York: Guilford Press.

Spearman, C. E. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Stemler, S. E., Grigorenko, E. L., Jarvin, L., & Sternberg, R. J. (2006). Using the theory of successful intelligence as a basis for augmenting AP exams in psychology and statistics. *Contemporary Educational Psychology, 31*(2), 344–376.

Sternberg, R. J. (Ed.). (1990). *Wisdom: Its nature, origins, and development.* New York: Cambridge University Press.

Sternberg, R. J. (1996). For whom does the Bell Curve toll? It tolls for you. *Journal of Quality Learning, 6*(1), 9–27.

Sternberg, R. J. (1997). *Successful intelligence.* New York: Plume.

Sternberg, R. J. (1998a). Abilities are forms of developing expertise. *Educational Researcher, 27*(3), 11–20.

Sternberg, R. J. (1998b). A balance theory of wisdom. *Review of General Psychology, 2,* 347–365.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology, 3,* 292–316.

Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist, 36*(4), 227–245.

Sternberg, R. J. (2002). Smart people are not stupid, but they sure can be foolish: The imbalance theory of foolishness. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 232–242). New Haven, CT: Yale University Press.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized.* New York: Cambridge University Press.

Sternberg, R. J. (2005). WICS: A model of leadership. *The Psychologist-Manager Journal, 8*(1), 29–43.

Sternberg, R. J. (2007). Finding students who are wise, practical, and creative. *The Chronicle of Higher Education, 53*(44), B11.

Sternberg, R. J. (2008a). Assessing what matters. *Educational Leadership, 65*(4), 20–26.

Sternberg, R. J. (2008b). The WICS approach to leadership: Stories of leadership and the structures and processes that support them. *The Leadership Quarterly, 19*(3), 360–371.

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence?* Norwood, NJ: Ablex.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., et al. (2000). *Practical intelligence in everyday life.* New York: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (1999). Myths in psychology and education regarding the gene-environment debate. *Teachers College Record, 100,* 536–553.

Sternberg, R. J., & Grigorenko, E. L. (2004). WICS: A model for selecting students for nationally competitive scholarships. In A. S. Ilchman, W. F. Ilchman, and M. H. Tolar (Eds.), *The lucky few and the worthy many. Scholarship competitions and the world's future leaders* (pp. 32–61). Bloomington, IN: Indiana University Press.

Sternberg, R. J., & Grigorenko, E. L. (2007). *Teaching for successful intelligence* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2009). *Teaching for wisdom, intelligence, creativity, and success.* Thousand Oaks, CA: Corwin.

Sternberg, R. J., & Jordan, J. (Eds.). (2005). *Handbook of wisdom: Psychological perspectives.* New York: Cambridge University Press.

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., Nokes, K., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., et al. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence, 29,* 401–418.

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., Reznitskaya, A., & Jarvin, L. (2007). Teaching for wisdom: What matters is not just what students know, but how they use it. *The London Review of Education, 5*(2), 143–158.

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., & Williams, W. M. (1996). *How to develop student creativity.* Alexandria, VA: Association for Supervision and Curriculum Development.

Terman, L. M. (1950). *Concept Mastery Test.* New York: Psychological Corporation.

Torrance, E. P. (1962). *Guiding creative talent.* Englewood Cliffs, NJ: Prentice Hall.

Torrance, E. P. (1974). *Torrance tests of creative thinking.* Lexington, MA: Personnel Press.

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology, 52,* 1236–1247.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology, 49,* 436–458.

Wagner, R. K., & Sternberg, R. J. (1990). Street smarts. In K. E. Clark & M. B. Clark (Eds.), *Measures of leadership* (pp. 493–504). West Orange, NJ: Leadership Library of America.

Wagner, R. K., & Sternberg, R. J. (1991). *Tacit knowledge inventory for managers.* San Antonio, TX: Psychological Corporation.

Wagner, R. K., Sujan, H., Sujan, M., Rashotte, C. A., & Sternberg, R. J. (1999). Tacit knowledge in sales. In R. J. Sternberg & J. A. Horvath (Eds.), *Tacit knowledge in professional practice* (pp. 155–182). Mahwah, NJ: Erlbaum.

Wallach, M., & Kogan, N. (1965). *Modes of thinking in young children.* New York: Holt, Rinehart, & Winston.

Yamamoto, K. (1964). Creativity and sociometric choice among adolescents. *Journal of Social Psychology, 64,* 249–261.

16 The Making of a Discipline

DAVID R. OLSON

In her authoritative review of the history of educational research, Ellen Lagemann (2000) asked why a paradigmatic educational theory sustained by systematic research had failed to develop over the past 100 years in spite of the rich diversity of research devoted to education. On the one hand, almost everything imaginable is, in some sense, relevant to education, hence the diversity of research; and on the other hand, the focus on raising test scores by whatever means, the legacy of Thorndike's behaviorism, placed a rigid constraint on originality and diversity. Thus, Lagemann is able to explain the almost universal enthusiasm for the progressive, liberal educational philosophy of John Dewey and the almost total disregard for such enlightened views in modern bureaucratic forms of schooling. Lagemann held some hope for a more sustained impact from the developing cognitive sciences, and the chapters in this book deliver, to some extent, on that promise.

Thorndike identified teaching with learning—no learning, no teaching—and attempted to set out general laws of learning—repetition, effect, relatedness—and predicted learning effects in terms of those laws together with the measurement of the abilities that the learner brought to the learning situation. As Schoenfeld (this volume) notes, current research goes beyond general theories of learning by addressing learning in domain-specific contexts with domain-specific content knowledge and domain-relevant epistemologies—that is, ways of thinking. And as Stanovich and Stanovich (this volume), Kaufman, Beghetto and Baer (this volume), Kozulin (this volume), and Sternberg (this volume) show, current research goes beyond using IQ as a generic predictor to the more nuanced study of what it means to be rational and competent in particular content-relevant domains.

Human development is a far broader topic than is education and is the topic of a broad range of sciences ranging all the way back to the study of the genome, as exemplified by Grigorenko, Mandelman, Naples, and Rakhlin (this volume). Education, on the other hand, is a more deliberate, goal-directed activity and, as my concern is with the development of a disciplined study of education, I shall comment primarily on education as an institutional practice in a modern society. Like other forms of human activity, education is the enterprise of formulating a set of goals—roughly, desires—and fulfilling them in the light of socially validated beliefs through a series of intentional practical actions. And to see that the goals have been achieved, one checks for the alignment of the goals with the outcomes—hence, in-school testing.

Schoenfeld (this volume) set out this scheme in its general form and applied it to teaching mathematics. That scheme is a descendant of the earlier Miller, Galanter, and Pribrim (1963) TOTE model, in which action is described as follows: T, test to see if some valued goal is met; O, operate on the available representations; T, test again now to see if the goal is met; and if so, E, exit. If not, return to another O, operate on the available representations. In the early 1960s, the model was rejected by some psychologists because they were reluctant to admit so cognitive a state as a goal or an intention, let alone representations or beliefs on which one could operate. That reluctance is now gone and Schoenfeld can describe teaching in terms of setting an agenda, including the goals to be achieved by the learner, eliciting responses from the learners, deciding whether or not to accept the responses and move on or elicit further comments and discussion, and repeating the cycle until, in the teacher's judgment, the goals have been achieved. Further confirmation comes, of course, with assignments and testing. Although algorithmic, this is far from a mechanical procedure and, as Schoenfeld points out, it is guided at every step of the operation by the beliefs, goals, and intentions of the teacher and students. Teachers may attempt to honor two or more goals at the same time and be forced to choose between inviting discussion to encourage students to express their own views and terminating discussion in order to move on to more important issues. Thus, in order to understand any teacher's performance, one must understand his or her knowledge about the domain and the goals of the activity as well as his or her beliefs about the student's confusions, and views regarding appropriate teaching and learning processes. Other writers add only that at least some of these decisions may be turned over to the learners themselves.

Nussbaum, Gomez, Mena, Imbarack, Torres, Singer, and Mora (this volume) have implemented, to good effect, a version of this TOTE model in a computer-supported collaborative learning (CSCL) environment. The technology enables the learners both to think on their own and to learn in groups while taking individual responsibility for their activities. The activities are problem-centered and the pedagogy is based on giving and evaluating reasons for any opinion. It further requires learners to come to an agreed-upon conclusion. Teachers monitor the progress of groups with their own computers and intervene when group discourse breaks down. The advantage of the technology is the evidence it provides of the activities of the learners; that all learners have computers on which they are to create their own conclusions after discussion with the group, puts an important focus on the agency and responsibility of individuals for their own learning as well as for the functioning of the group.

While teaching is the heart of the educational process, it is far from the only relevant practice. Schooling is also expected to impart a sense of orderliness and an understanding of and commitment to a set of norms and procedures for attending, participating, and respecting the teacher and other students. Indeed, teaching as described above presupposes that students already share the "format" of the lessons—namely, the right and responsibility of the teacher to set the agenda, the obligation to respond to the teacher's directives, to make their comments relevant to the topic at hand, and to adopt the goal set out by the teacher. "Class dismissed" is the teacher's prerogative, as students quickly learn; other norms, largely those of a Gricean form—be relevant, be perspicuous, be brief—are more difficult to learn and to enforce. As Rothstein-Fisch, Greenfield, Trumbull, Keller, and Quiroz (this volume) point out, Latino children in particular, but many children in multicultural contexts, bring cultural norms that are more communal and cooperative, whereas the norms of the traditional classroom are more individualist and competitive. These authors provide evidence that schools can be reoriented to acknowledge and accommodate such differences with good effect both on the self-respect of the learners and on their school achievement. Lin, Siegler, and Sullivan (this volume) note the importance of the learner's goals, which in some cases may be negotiated to mutual advantage with the teacher. These goals determine what learners do and learn. Consequently, contra Schoenfeld, it is not simply a matter of the teacher announcing an agenda but rather the fraught matter of teacher and learner coming to some mutually acceptable set of goals and criteria for learning and acting for which both may be held accountable.

Perhaps the most original proposal discussed in this volume, widely endorsed and in some cases empirically validated, involves what in the Schoenfeld decision model was the follow-up to a learner's response: does the teacher take the further step of exploring the learner's misunderstandings or does the teacher simply move on to the next idea? Nuñes and Bryant (this volume) take the important step of having the teacher recognize that exploring a learner's reasons for an incorrect answer is just as valuable as exploring the reasons for advancing the correct answer. They point out that, in both cases, children have reasons, often implicit, for any suggestions they make and any answers they offer. Discussing and evaluating these reasons is often critical for getting learners to revise their beliefs. Again, of course, the teacher has to make a judgment as to whether and when to cut losses.

In the same vein, that is, on the elaboration of classroom procedures for getting children to give and share reasons for their answers, is the proposal by Resnick, Michaels, and O'Connor (this volume) that the practice of giving of reasons is best seen as a social process. That is, it is important that learners "give reasons for your answers," a well-worn pedagogical move, but even more important that they advance those reasons to other learners, who can then criticize and build upon them. In fact, there is a growing commitment to group work as opposed only to solitary "seat" work. As those authors note, there is group work and there is "group work." What is critical is that a culture of discussion develops—of the students themselves giving and evaluating their reasons for believing as well as their reasons for doubting. As Resnick and colleagues point out, for reasons to have an impact on thinking, it is essential that they be clearly formulated and that the norms for evaluating them be explicit and shared. Such discourse, of course, is riddled with metacognitive talk about what is said and what is meant as well as the reasons for saying it and the validity of so saying. Putting the obligation on to the learners has as a side benefit—the fact that the responsibility for learning falls increasing on the learner—a point to which I shall return.

Contrary to the "folk psychological" notion that the teacher is the one with knowledge and the learners are the ones who are ignorant, requiring only to be told, these modern notions see the learner quite differently. They set out to discover the intuitions and prejudices that learners bring to the task and they use these discoveries as the basis for managing the discourse of the group. Nuñes and Bryant (this volume) show, for example, that although children have difficulty grasping the mathematical notion of proportions, they have strong, presumably universally held intuitions about one-to-many mapping. These intuitions, once made explicit, may be translated, perhaps with some difficulty, into more abstract mathematical notions.

Not everyone immediately sees the importance of elaborating reasons. Lin, Siegler, and Sullivan (this volume) found that many students simply ignore incorrect answers, as well as unexpected statements and events, rather than looking for possible explanations. It may require a teacher's initiative to get learners to look for disconfirming evidence, and it may require expert teaching skill to get learners to see that the answers, known to be false, may in fact be rational, or grounded in reasons. Those reasons, in turn, can then be examined in terms of evidence and logic as well as for their possible promise.

In valuing the learner's reasons and the activity of explicating and evaluating them, indeed in seeing learners' actions and answers as rational even when they are wrong, I am making a small dispute with Stanovich and Stanovich (this volume) about the use of the term *rational*. I do so while endorsing their main claim that IQ is at best a crude index of the ability to make reasoned and informed judgments and that the latter is an important responsibility of the school. The reasons and reasoning advocated by Lin and colleagues, Resnick and coworkers, Nuñes and Bryant, Schoenfeld, Kozulin, and others hew close to the conventional meaning of *rationality*, namely, justified by reasons. In fact, Stanovich and Stanovich devote much of their research to discovering the reasons thinkers use in leading themselves to unwarranted conclusions. These reasons are often bad ones and, when recognized as such, may be held in check. They include anchoring bias, confirmation bias, hindsight bias, self-serving bias, base-rate neglect, and other failings to which "this flesh is heir." But the answers are typically rational—grounded in reasons—but nevertheless not valid or just not true. The same is true of children's reasoning; it is not irrational but rather grounded in unshared premises, which, when made explicit, are found to be rational if misleading or inappropriate to the immediate context. In fact, teaching thinkers to be more rational involves the very activity of giving and evaluating ideas through discussion with others of both correct and incorrect answers.

It also involves learning a metalanguage for managing such epistemological discussion. These principles are recognized both by Stanovich and Stanovich and by the other researchers who emphasize the rational discourse discussed above.

The overwhelming endorsement of content-based, group-organized, reasoning-based pedagogy found in this volume must be seen in relation to the actual, average pedagogy more generally found in the classrooms of the world. Preiss (this volume) found little evidence in the Chilean schools he studied of the giving and evaluating of reasons by the learners themselves. Rather, classroom talk was largely teacher talk, consisting of conveying information and controlling the actions of students. There was little emphasis on reflection or elaboration and little on the expression of opinion or the use of metalinguistic and metacognitive language for guiding the talk on behalf of the teachers, let alone the students. Preiss traces this lack to the "folk pedagogy" of the teachers—their beliefs about knowledge, the learner, and the appropriate actions for a teacher to take. Changing teachers' behavior involves not just learning some new skills but also becoming aware of their pedagogical assumptions so that they may see the possibility of new ones. Without a change in their folk pedagogy, new methods will simply be reduced to the ancient ritual of telling and testing. Nor is this only a Chilean problem. Preiss examined the video lessons filmed around the world by the TIMMS project and found that none of the lessons illustrated "reasoned methods of verification"—that is, using reasons and reasoning as a primary means of learning and arriving at rationally justified beliefs.

One place where the giving and evaluating of reasons as a pedagogical strategy continues to be of marginal relevance if not ignored is in the study of learning to read. Wagner and Kantor (this volume) point out that normal reading requires "adequate vision; eye-movement control; visual perception; knowledge of the orthographic, phonological, and semantic features of words; higher-level linguistic and cognitive skills; memory; and attention." A primary bottleneck for all beginning readers, especially those attempting to master an "opaque" writing system such as the English alphabet and especially for dyslexic children, is the ability to segment the stream of speech into the subsyllabic constituents that are represented by letters of the alphabet. But unlike learning in the other domains considered in this volume, reading ability seems to be unaffected by the giving and examination of learners' reasons. The people who created and those who revised the alphabet clearly had reasons for their actions, but these reasons seem largely incomprehensible (e.g., "when two vowels go a-walking the first one does the talking") and are of little use to readers. Or perhaps they need to be further explored. In any case, dyslexic children do have difficulties in reading and spelling, although, at least in Canada, almost all children are able to read and write to at least a limited extent by the time they are 10 or 12. Whether authorities find that an acceptable standard is a much-debated question. Although I, of all people, do not need to be persuaded of the importance of writing, it is my view that the overwhelming attention currently being paid to literacy in the primary school should be tempered by an increased focus on content and on oral discourse, with writing only one of several media for gaining and expressing ideas.

One procedure that may go some distance in turning learning to read into a more rational process is what, in Cynthia McCallister's "genre practice" program (www.genrepractice.org), is called "unison reading." Children, even at the earliest stages of learning to read, read orally together, and when a mismatch occurs, the process stops and children discuss among themselves what went wrong, sometimes with the teacher intervening when the discussion breaks down. In this practice, discussion does often turn metalinguistic, while it is also rational and communal. It remains to be seen if this pedagogy is as effective as it promises to be.

Here I endorse the ambitious and forward-looking activities developed by Zigler and Finn-Stevenson for their School of the 21st Century. Recognizing that schooling alone is inadequate to the education of an increasingly large number of children, they have created a program that combines child care from an early age with health services and family services and coordinated it with the ongoing activities of the school. While day care and preschool programs are spotty, diverse, and of varying quality, Zigler and Finn-Stevenson designed and have now implemented in some 1,300 schools a set of activities and resources constituting an informed and effective program centered in the school. For the program to succeed, they point out the importance of having on board all of the agents of the school, from the boards of education to the principals to the practicing teachers. Importantly, they move away from the view that preschool education is just primary education started earlier to the view that there are many fundamental things to be learned—including self-reliance, cooperation, an interest in explanation, the nature of rules, the variety of roles, and the like—that are far more important than learning the ABCs.

Intelligence continues to dominate much of the discourse of psychology and education. Few schools set as a goal the enhancement of intelligence, although I see no reason why it should not be, especially if we see intelligence as mental age rather than as individual difference, that is, in IQ terms. The nature of intelligence continues to be debated. Much recent research, as mentioned at the outset, shows that much more than IQ is involved in what Gardner, Csikszentmihalyi, and Damon (2001) have called "good work." Sternberg (this volume) analyzes what he calls "successful intelligence," what one needs to be intelligent, informed, and wise, and he attempts to develop a test, comparable to the SAT, that would better predict not only who learns well but who does well. Stanovich and Stanovich (this volume) show how poorly IQ predicts the ability to make high-level rational decisions and call for programs that would combat the common failings of reasoners mentioned earlier. Kozulin (this volume), too, wants to go beyond measures of IQ to test for what students are capable of learning with help, what he calls "dynamic assessment," to augment traditional measures.

The importance of these proposals, as I see it, is in what they offer, not so much to satisfy our addiction to testing as in helping us in reenvisioning the goals and practices of education. For more than a century, educational theorists have been urging a greater focus on *doing* and on *process,* on *thinking* rather than simply on *learning.* The extension of the notion of intelligence into practical judgments, rationality, and the exploration of ideas with others should not be left aside while the school relentlessly attacks basic skills and testable competencies. School should be an environment for the advance and discussion of ideas, of rational discourse, and of taking on and meeting the standards of progressive, rational argument. All of the proposals for going beyond IQ discussed above make concrete suggestions as to how intellectual competence could be enhanced through such discourse (Wells, 1999).

The entire discourse of traits and dispositions, of which intelligence is sometimes thought of as one, however, falls foul of the philosophical argument that traits and dispositions "lack normative force." In other words, dispositions are things for which we are not responsible and the exercise of which are largely inappropriate as a basis for earning approval and other forms of credits. Character, on the other hand, describes properties for which one is responsible and for which one may take some credit. Education is a matter of helping people become responsible for their beliefs and actions, by which one thereby earns entitlements. Entitlements, including privileges and rewards, are earned by responsibly meeting valued goals, and more generally by becoming responsible for who we are (for further discussion, see Olson, 2006).

Piaget claimed in the title of one of his many books that "to invent is to understand," a notion that provided a basis for discovery learning. Learning is not something that someone else can do for you; it is something you must do for yourself. As Renninger (this volume) points out, it both calls for and develops the learner's self-efficacy and self-regulation. When the learner has a role in the setting of goals, he or she also takes on responsibility for the achievement of those goals. Achieving them adds to confidence and a willingness to take on even more challenging goals. Pedagogical reforms that emphasize the giving and evaluating of reasons by the learners themselves provide just the environment not only for taking on goals but also for thinking about thinking—for rational discourse. Students working in a group, when focused on a problem, have been shown to raise possible solutions, understand both why those solutions may be wrong, and also understand why the person offering the solutions may have thought them to be correct. This is thinking about thinking in action, an important aspect of epistemological development. As all the above authors acknowledge, working in groups is more than sharing a table. It requires learning the rules and norms for discussion, for respecting the offerings of others, for giving everyone a chance, and for recognizing the importance of the validity of an argument. It also involves recognizing the importance of counterexamples and all of the other types of reasoning and reasoning errors discussed by Stanovich and Stanovich, as well as their appropriateness in context, as discussed by Sternberg.

THE MAKING OF A DISCIPLINE

I believe we have the ingredients for a discipline of education. Education in a modern society is an institutional practice comparable to the health care system, the financial system, the justice system, the government, and the academy. These systems take on and manage responsibility for the achievement of certain goals through the assignment of professional responsibilities and the allocation of resources. A discipline of education is the knowledge base and the set of rules for application of that knowledge base to the more specialized professional roles assigned in that institutional system—board member, principal, teacher, psychologist, and so on.

The very nature of these systems has resulted in a remarkable uniformity of educational practices worldwide. Neoinstitutionalist theorists LeTendre, Baker, Akiba, Goesling, and Wiseman (2001) showed that the international comparisons of schooling, such as those of the Third International Math-Science Study (TIMSS), were less informative than hoped because of the "institutional isomorphism" of school systems—the fact that all schools in modern bureaucratic societies are more or less the same in regard to mandated outcomes. Local culture tends to be swamped by the bureaucratic "rationalization of society and the spread of institutions that has taken place over the last century" (p. 12). It is a remarkable fact that cultures once seen as amazingly diverse are now so similar in this regard that a single metric, a reading test, can be applied worldwide with results not only interpretable but, in fact, very similar. Individual differences persist but group differences are small and tend to reflect the years of schooling. It is bizarre that a minor 2- or 3-point spread can set entire nations into a frenzy over "school failure" and lead to calls further to standardize and regiment the entire system! As Berliner and Biddle (1995) have argued, the differences in scores are more attributable to general demographic facts, primarily poverty and access to opportunity, than they are to explicit pedagogy.

The further downside of this bureaucratic model is the temptation to regard the educational system in terms of command and control, whereby the top levels of the bureaucracy determine what goes on at the lower levels, all the way down to the learner. But this is less a downside than a seriously misleading flaw when it comes to learning. The agency, I believe, is all wrong.

I have argued (Olson, 2003, 2006), as do the authors in the current volume, that reform in education is needed to make the learner "the active agent in their own learning process" (Nussbaum et al., this volume). It is the learner who does the learning. Teachers' goals are relevant only to the extent that they become the learner's goals. Put in another way, teachers' goals are to be found in the interests and predispositions of the learners. Learners gain satisfaction as well as a variety of other entitlements, ranging from praise and glory to the right to undertake even greater tasks and responsibilities, by succeeding in their learning and in taking on and meeting their responsibilities. The teachers' role is to meet their obligations both to the institutions that employ them and to the learners they serve. Their obligation to the learners is to help them formulate and take on worthwhile goals and to then monitor their successes. In a word, the teacher's responsibility is to make it possible for the learners to meet their goals.

The principal's responsibility is to make it possible for the teachers to meet their goals, and so on up the bureaucratic chain. As in the Schoenfeld model, the teacher, as an agent of the state, sets the agenda, but the goals that are pursued within that agenda must be negotiated with the learner so that the learner's goals are an acceptable compromise with the somewhat more lofty goals entertained by the school. Once goals are agreed upon, learners can accept responsibility for their own learning and the teacher's job shifts to helping them define acceptable criteria for successfully meeting them. And so on up the institutional chain.

Understanding the educational process in this way puts a new emphasis on the professional responsibilities of the teacher. Bureaucratic reform over the past two decades has tended to proceed by increasing attempts to regulate and thereby control the actions and decisions of teachers by providing closely scripted procedures selected by large-scale studies of "what works" and enforced by repeated testing. However, if teachers are to make the nuanced judgments discussed above, they must also be given the freedom to choose the means most appropriate to the moment to moment beliefs and reasons entertained by the learners. Freedom to decide is an essential component of taking on responsibility. A highly trained, professional teaching force capable of exercising this freedom is a promising alternative to the regimen of scripted programs and relentless testing.

The thrust of this proposal is to reorient the theory of education away from an emphasis on social control, that is, by imposing demands on students and monitoring compliance. Rather, it is to move toward an emphasis on the competence of the learners, conceived in terms of their ability to take on and meet their goals and their obligations—and of seeing the teacher's role as that of helping learners formulate valued and achievable goals and providing the resources needed so that they can achieve them.

With success in achieving one's goals and meeting one's responsibilities comes feelings of self-efficacy and self-control. These goals and obligations begin with the child and with learners' generally taking responsibility for the beliefs and actions they explore in the process of learning. The emphases on responsibility and obligation turn educational theory from a purely cognitive theory into a moral theory as well, just where critics from Kant to Dewey believed it should be. With the works discussed herein, I believe we have taken an important step in that direction.

REFERENCES

Berliner, D. C., & Biddle, B. J. (1995). *The manufactured crisis.* New York: Addison-Wesley.

Gardner, H., Czikszent Mihalyi, M., & Damon, W. (2001) *Good work: Where excellence and ethics meet*. New York: Basic Books.

Le Tendre, G., Baker, D., Akiba, M., Goesling, B., & Wiseman, A. (2001). Teacher's work: Institutional isomorphism and cultural variation in the U.S., Germany, and Japan. *Educational Researcher*, 30, 3–15.

Lagemann, E. C. (2000). *The elusive science: The troubling history of education research.* Chicago: University of Chicago Press.

Miller, G., Galanter, E., & Pribrim, K. (1963). *Plans and the structure of behavior.* New York: Holt, Rinehart and Winston.

Olson, D. R. (2003). *Psychological theory and educational reform: How schools remake mind and society.* Cambridge, UK: Cambridge University Press.

Olson, D. R. (2006). Becoming responsible for who we are: The trouble with traits. In *Howard Gardner under fire: A rebel psychologist faces his critics* (pp. 39–44). Chicago: Open Court.

Wells, G. (1999). *Dialogic inquiry: Towards a socio cultural practice and theory of education*. New York: Cambridge University Press.

Index

This page intentionally left blank